PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12Q 1/68, C12P 19/34, 21/08, G01N 33/53, 33/567, 33/574, C07H 21/04, C07K 1/00

(11) International Publication Number: A1

WO 96/29433

(43) International Publication Date: .26 September 1996 (26.09.96)

(21) International Application Number:

PCT/US96/03940

(22) International Filing Date:

22 March 1996 (22.03.96)

(30) Priority Data:

08/409,823

23 March 1995 (23.03.95)

(71) Applicant: RESEARCH FOUNDATION OF STATE UNI-VERSITY OF NEW YORK [US/US]; The W-5510 Melville Library, Office of Research Services Stony Brook, Stony Brook, NY 11794-3366 (US).

(72) Inventors: MANDEL, Gail; 7 Emmet Way, Stony Brook, NY 11790 (US). CHONG, Jayhong, Andrew; 9 Longhorn Lane, East Satauket, NY 11733 (US).

(74) Agent: BLOOM, Allen; Dechert Price & Rhoads, Princeton Pike Corporate Center, P.O. Box 5218, Princeton, NJ 08543-5218 (US).

(81) Designated States: AL, AM, AU, BB, BG, BR, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KP, KR, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: REST PROTEIN AND DNA

(57) Abstract

The invention provides a substantially pure nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues. The invention also provides a substantially pure nucleic acid encoding a protein that binds to a promoter sequence having at least about 90 % homology to nucleotides 6-28 of the RE1 sequence and acting to suppress the acitivity of a promoter having the promoter sequence. The invention further provides a substantially pure nucleic acid encoding a protein having at least about 85 % homology to at least one of the DNA binding domain or the suppressor domain of an animal RE1-Silencing Transcription factor. The invention also relates to the proteins so encoded.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
ΑÜ	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Кевуа	RO	Romania
BY	Beiarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland :	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	L	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Larvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

10

15

20

- 1 -

REST PROTEIN AND DNA

The present invention is directed to purified nucleic acids encoding RE1-Silencing Transcription factors ("REST proteins") and to purified proteins with REST activity.

Part of the work performed during the development of this invention utilized United States Government Funds under National Institutes of Health Grant NS22518 and National Science Foundation Grant GER9023237. The government has certain rights in this invention.

It has been suggested that neural development is substantially a default pathway of development that is repressed in non-neural cell types. Consistent with this idea, Kraner et al., Neuron 9, 37-44, 1992, identified a DNA sequence, the 28 base pair ("bp") RE1 sequence, found in the 5' flanking sequence of the gene for the membrane protein that forms the CNS-type voltage dependent sodium channel (i.e., "type II" voltage dependent sodium channel), that appears to be responsible for negatively regulating the use of this gene in non-neural tissue. RE1 nucleic acid sequences also appear to interact with a nuclear protein found in non-neural calls but not in most neural cells. Similar sequences having cell-specific silencer activity have been identified in the promoters for SCG10 (Mori et al., Neuron 9, 45-54, 1992), synapsin (Li et al., Proc. Natl. Acad. Sci. USA 90, 1460-1464, 1993) and dopamine β -hydroxylase (Ishigoro et al., J. Biol. Chem. 268, 17987-17994, 1993).

Summary of the Invention

Until now, however, the protein responsible for silencing promoters containing RE1 elements has not been identified. That protein herein referred to as "REST," and the gene encoding it, is herein identified as having the amino acid sequence included in SEQ ID NO:1. The portion of the nucleic acid sequence included in SEQ ID NO:1 that is an open reading frame for REST is identified as SEQ ID NO:10. The protein sequence for human REST and the nucleic acid sequence of the CDNA for human REST are shown in Figure 1.

One preferred embodiment of the present invention is a substantially pure nucleic acid

comprising a nucleic acid encoding a protein having at least about 85% homology to at least the

DNA binding domain or the suppressor domain of an animal REST protein; the same substantially

pure nucleic acid further comprising a nucleic acid encoding at least the DNA binding domain or

the suppressor domain of an animal REST protein; the same substantially pure nucleic acid,

wherein the REST protein is a mammalian REST protein; the same substantially pure nucleic acid,

wherein the REST protein is a human REST protein; the same substantially pure nucleic acid,

10

15

20

25

30

wherein the nucleic acid comprises SEQ ID NO:2; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:10; the same substantially pure nucleic acid, further comprising a nucleic acid encoding both the DNA binding domain and the suppressor domain of an animal REST protein; the same substantially pure nucleic acid, wherein the REST protein is a mammalian REST protein; the same substantially pure nucleic acid, wherein the REST protein is a human REST protein; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:2; the same substantially pure nucleic acid, wherein the nucleic acid comprises SEQ ID NO:10; the same substantially pure nucleic acid, comprising a nucleic acid encoding a protein differing from an animal REST protein by no more than about 20 point mutations. Preferred substantially pure nucleic acids also encode analogs to the REST protein, which include either the DNA binding domain or the suppressor domain thereof.

Another preferred embodiment of the present invention is a substantially pure nucleic acid that hybridizes with an animal REST nucleic acid under stringent conditions; the same substantially pure nucleic acid, comprising the nucleic acid of SEQ ID NO:1.

A further preferred embodiment is a substantially pure nucleic acid comprising a nucleic acid encoding a protein that binds to a promoter having at least about 90% homology to nucleotides 6-28 of SEQ ID NO:29 and acting to suppress the activity of a promoter having said promoter.

Yet another preferred embodiment is a substantially pure protein having at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure protein, comprising at least the DNA binding domain or the suppressor domain of an animal REST protein; the same substantially pure protein, further comprising the protein of SEQ ID NO:2; the same substantially pure protein, further comprising both the DNA binding domain and the suppressor domain of an animal REST protein; the same substantially pure protein, further comprising the protein of SEQ ID NO:10.

Yet another preferred embodiment is a transformed eukaryotic or prokaryotic cell comprising a nucleic acid encoding a protein having at least about 85% homology to at least one of the DNA binding domain or the suppressor domain of an animal REST protein; the same transformed cell, further comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain of an animal REST protein; the same transformed cell, wherein the REST protein is a mammalian REST protein; the same transformed cell, wherein the REST protein is a human REST protein; the same transformed cell, wherein the nucleic acid comprises SEQ ID NO:

2. Preferably, the transformed cell expresses one of the inventive proteins described herein.

Yet another preferred embodiment is a vector capable of reproducing in a eukaryotic or prokaryotic cell comprising a nucleic acid encoding a protein having at least about 85% homology to at least the DNA binding domain or the suppressor domain of an animal REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, further comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain of an animal REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the REST protein is a mammalian REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the REST protein is a human REST protein; the same vector capable of reproducing in a eukaryotic or prokaryotic cell, wherein the nucleic acid comprises SEQ ID NO:2. Preferably, the inventive vector expresses, intracellularly or extracellularly, one of the inventive proteins described herein.

- Yet another preferred embodiment is a method of preparing a protein having REST activity, wherein the protein has at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein, the method comprising:
- (a) transforming an appropriate eukaryotic or prokaryotic cell with an expression vector for expressing intracellularly or extracellularly a nucleic acid encoding the protein;
- 15 (b) growing the transformed cell in culture; and
 - (c) isolating the protein from the transformed cell or the culture medium.

Yet another preferred embodiment is a pharmaceutical composition for treating an animal having de-differentiated neural cells or neural cells exhibiting diminished activity comprising an effective amount of a REST-interfering nucleic acid, wherein the REST-interfering nucleic acid comprises an antisense molecule directed against REST expression or an expression vector for expressing REST DNA binding activity but not REST silencer activity, and a pharmaceutically acceptable carrier; the same pharmaceutical composition, wherein the animal has brain cancer; the same pharmaceutical composition, wherein said animal has a demyelinating myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies, traumatic nerve injury, post stroke degeneration, post-traumatic spinal and neural degeneration, poliomyelitis or rabies.

Yet another preferred embodiment is a pharmaceutical composition for an animal having neural cells exhibiting excessive neural activity comprising an effective amount of an expression vector comprising a nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues, and a pharmaceutically acceptable carrier; the same pharmaceutical composition, waterein the animal has epilepsy, Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke or a neurodegenerative disease; the same pharmaceutical composition,

wherein the animal has Alzheimer's, Parkinson's or Huntington's disease; the same pharmaceutical composition, wherein the animal has epilepsy; the same pharmaceutical composition, wherein the animal has a neurodegenerative disease.

Yet another preferred embodiment is a method of determining the level of REST expression in 5 tissue sample comprising

- (a) contacting the tissue sample with (i) a nucleic acid that binds to REST mRNA under stringent conditions or (ii) an antibody specific for REST;
- (b) washing the tissue sample to remove non-specific hybridizations of the nucleic acid or non-specific antibody binding; and
- 10 (c) determining the level of hybridized nucleic acid or bound antibody.

Yet another preferred embodiment is an antibody that reacts specifically with the substantially pure protein having at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein, as recited above.

Breef Description of the Drawings

Figure 1 shows the protein encoded by the open reading frame of SEQ ID NO:1 and the nucleotide sequence of SEQ ID NO:1.

Detailed Description of the Invention

The DNA binding domain of REST is made up of eight zinc finger domains. The portion of SEQ ID NO:1 that encompasses the eight zinc finger domains of REST is identified as SEQ ID NO:2. The underlined residues shown in Figure 1 are the zinc finger domains. A search of the GenBank database found that the closest homology for this DNA binding domain is found with the Krüppel family of repressor proteins, particularly the GLI-Krüppel repressor protein. (For a review of zinc finger proteins, see Colman, Ann. Rev. Biochem. 61, 897-946, 1992.) The size of the RE1 sequence, 28 bp, and the number of zinc finder domains in REST is consistent with research (Pauletich and Pabo, Science 242, 809-817, 1991) that suggests that each such zinc finger domain interacts with a triplet of nucleotide base pairs.

The sequences of the zinc finger domains are indicated in the table below (with a space inserted into 6 of the 8 sequences to facilitate alignment of homologous sequence):

SEQ ED NO.	Zinc Finger Sequence
11	CKPCQYEAESEEQFVHHIRV H
12	CDRCGYNTNRYDHYTAHLKH H
13	CIICTYTTVSEYHW RKHLRN H
14	CGKCNYFSDRKNNYVQHVRT H
10 15	CELCPYSSSQKTHLTRHM RT H
16	CDQCSYVASNQHEVTRHARQVH
17	CPHCDYKTADRSNFKKHVEL H
18	CPVCDYAASKKCNLQYHFKSKH

15 C-terminal to the DNA binding domain, REST has six repeat sequences having the following sequences:

SEQ ID NO.	Internal Homologous Sequences
20 21	M E V V Q E G P A Q K E L L P P
22	M Q V V Q K E P V Q M E L S P P
23	MEVVQKEPVQIELSPP
24	MEVVQKEPVKIELSPP
25	I EVVQKEPVQM ELSPP
25 26	M G V V Q K E P A Q R E P P P P

These sequences are indicated in Figure 1 by the double underlined amino acid residues. The sequence encompassing these repeats is designed SEQ ID NO:20. The most highly conserved residues of the six repeats are highlighted in the table above.

By studying the activity of the RE1 promoter, it has been determined that REST is expressed in undifferentiated neural progenitors, which is consistent with the view that REST plays a role in maintaining the undifferentiated state of these cells. Antisense oligonucleotides directed

against the REST transcript accordingly, would promote the differentiated state. Also consistent with this view is the hypothesis that certain neuroblastoma cells have de-differentiated into analogs of neural progenitors. Accordingly, REST antisense therapy aides in reversing this de-differentiation and reducing or reversing the malignancy of these cells.

- As used herein, a "REST nucleic acid" means the REST-encoding nucleic acid, whether RNA or DNA, synthetic or natural, found in a REST-expressing animal, or the complementary strand thereof. "REST protein-encoding nucleic acid" or "nucleic acid encoding a REST protein" refers to any nucleic acid, whether native or synthetic, RNA, DNA, or cDNA, that encodes a REST protein. For recombinant expression purposes, codon usage preferences for the organism in which such a nucleic acid is to be expressed are advantageously considered in designing a synthetic REST protein-encoding nucleic acid. A "REST protein" is a REST homologous protein with the ability to bind an RE1 sequence and to repress the activity of a promoter containing an RE1 sequence. An "animal REST protein" is a REST protein expressed by a member of the animal kingdom; a "human REST protein" is a REST protein expressed by a human.
- Vectors encoding a protein with RE1-binding activity but not suppressor activity are shown herein to reverse the transcriptional suppression caused by REST, apparently by competing for the RE1 promoter element through which REST functions. Accordingly, gene therapy with such vectors are used like the aforementioned and other antisense therapies known in the art to reduce REST's suppressor activity. The vectors described in this paragraph and the antisense molecules decused above are termed herein "REST-interfering nucleic acids."

Probes for REST expression are used to measure the extent of a de-differentiation in biopsy tissue from tumors that are derived from neural tissue. Such probes are used to predict the extent of tissue transformation and the virulence of the tumor. Such probes include antibodies directed against REST or fragments thereof, nucleic acid probes that hybridize to REST mRNA under specifically prime a PCR amplification of REST mRNA.

For a number of years physicians have sought to treat neurodegenerative diseases by administering neural stem cells, for instance stem cells derived from embryos, to produce replacements for a patient's lost neural cells. Such diseases include Alzheimer's disease, Packinson's disease, Huntington's disease, amyotrophic lateral sclerosis ("Lou Gehrig's disease") and demyelinating diseases such as multiple sclerosis. Stem cells used in these therapies are induced to initiate differentiation to provide the needed replacement cells by treating them with

REST antisense constructs or with vectors expressing the DNA-binding domain of REST but not the suppressor function of REST.

In diseases where pathological states are associated with excesses in neural activity, such as epilepsy, Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke and neurodegenerative diseases (including Alzheimer's, Parkinson's and Huntington's diseases), the level of neural expression of the voltage-dependent sodium channel is usefully reduced. Toward this end, neural cells are transformed to express sufficient REST to down-regulate expression of the sodium channel.

In diseases that exhibit insufficient neural activity, such as demyelinating diseases (including multiple sclerosis), myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies, traumatic nerve injury, post-stroke degeneration, post-traumatic spinal cord neural degeneration, poliomyelitis and rabies, up regulation of the expression of the neural voltage-dependent sodium channel is useful. This up regulation is done by antisense therapy based on REST nucleic acids to inhibit neural expression of REST or with gene therapy using a vector that expresses a protein that competes with REST for RE1 promoter sequences without suppressing the activity of the promoter.

The REST protein is also a useful target for drug screening efforts to identify drugs that interfere with its suppressor activity, either by inhibiting DNA binding or the negative effect of REST on transcription. Such drug screening assays in one embodiment include cell-free transcription systems using the REST protein, cell-free transcription systems such as those described by Dignam et al., Nucl. Acids. Res. 11, 1475-1489, 1983 or that described in the cell-free transcription protocol available from Promega (Madison, WI) in an appropriate RE1-containing promoter. The screening methods also utilize in other embodiments expression studies conducted in cell culture, such as the chloramphenicol acetyl transferase (CAT) assay methods described herein below.

The suppression domain of REST is fused by recombinant methods to a DNA-binding domain of a positive transcription factor to create a protein that represses the activity of one or more promoters. For instance, in one embodiment the suppressor domain is linked to pit-1, a transcription factor for the prolactin and growth hormone promoters (see Ingraham et al., Cell 55, 519-529, 1988), thereby creating a vector for gene therapeutics aimed at down regulating hyperactive pituitary production of growth hormone and/or prolactin. Other examples of specific targets for this kind of therapy are the DNA-binding domains of steroid hormone or thyroid hormone receptors. Fusion vectors expressing a DNA binding domain from a steroid hormone receptor and the REST suppressor domain are used in yet other embodiments to down regulate

responsiveness to the steroid hormones in patients that overproduce the steroid or that have steroid hormone receptors that are too active. The fusion protein in one embodiment includes the target DNA-binding element and substantially all of the REST protein.

The antibodies and nucleic acid probes of the present invention are also useful as histochemical reagents for marking the pathways of nerves that do not express the CNS-type sodium channel. Also, the staining of most non-neural tissue serves as a contrast agent to highlight neurons that do not express REST or express very low levels of REST. Thus, these histochemical agents are used to produce histochemical slides and preserved anatomy specimens useful for training students and physicians.

The first embodiment of the invention relates to a purified nucleic acid comprising a nucleic acid having at least 85% homology to at least the DNA binding domain or the suppressor domain of an animal REST protein. Such a nucleic acid is referred to herein as a REST protein that binds the RE1 promoter element and/or suppresses the activity of the promoter for the CNS-type voltage-dependent sodium channel. The encoded protein is preferably a REST protein of a mammalian animal, more preferably the human REST protein. Preferably, the encoded protein has the sequence of SEQ ID NO:1, SEQ ID NO:2, or SEQ ID NO:10.

Another embodiment of the invention provides for one or more nucleic acids encoding a protein that binds to a promoter sequence having at least about 90% homology, preferably 95% homology, to nucleotides 6-28 the RE1 sequence (SEQ ID NO:29) and acting to suppress the a20vity of a promoter containing that promoter sequence. Yet another embodiment provides for a nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural tissues.

The nucleic acid embodiments of the invention are preferably deoxyribonucleic acids, preferably double-stranded deoxyribonucleic acids, except that, for hybridization probes, single-stranded nucleic acids are preferred. However, nucleic acids of the present invention also include remoderacids. The nucleic acids of the present invention are also referred to as polynucleotides or polynucleic acids.

Numerous methods are known to delete a segment of a nucleic acid from or mutate a nucleic acid that encodes a protein and to confirm the function of the proteins encoded by these deleted or mutated nucleic acids. Accordingly, the invention also relates to a mutated or deleted vaction of a REST protein-encoding nucleic acid that encodes a protein that retains the ability to bind specifically to the RE1 promoter element and/or the ability to suppress an RE1-responsive promoter when appropriately bound to the vicinity of the promoter.

The invention also relates to a nucleic acid encoding, in the proper order, at least 4 of the zinc finger domains of a REST protein, preferably at least 6 of the zinc finger domains, more preferably all of the zinc finger domains. The zinc finger domains for human REST are identified in Figure 2. Preferably, the nucleic acid is SEQ ID NO:2.

Transcription suppressive proteins, such as Krüppel, Kid-1, and ZNF2 generally have distinct suppressor domains which function so long as they are appropriately linked to DNA binding domains that suitably bring the suppressor domains into the vicinity of the target promoters. See, for instance, Licht et al., Nature 346, 76-79, 1990; Witzgall et al., Proc. Natl. Acad. Sci. USA 91, 4514-4518, 1994. Such a suppressor domain can readily be identified for the REST protein using defectional approaches and recombinant fusion protein approaches that are well known in the art. Accordingly, the invention also is directed to a nucleic acid encoding a segment of the protein of a REST protein that is effective to repress the use of a promoter when attached to a protein that binds the promoter. Preferably, the encoded protein will be effective to repress the use of the promoter for the CNS-type voltage-dependent sodium channel gene. Studies with the aforementioned RE1 nhbleic acid suggest that it is ineffective as a transcription silencing element when inserted into some gene promoters. Accordingly, the promoters discussed in reference to this embodiment are RE1-responsive promoters.

It is recognized that many deletional or mutational analogs of nucleic acid sequences for a REST protein are effective hybridization probes for REST nucleic acid. Accordingly, the invention relates to nucleic acid sequences that hybridize with such REST-encoding sequences under stringent conditions. Preferably, the nucleic acid of the present invention hybridizes with SEQ ID NO:1 under stringent conditions. The invention also relates to nucleic acids that hybridize with SEQ ID NO:2 under such stringent conditions.

"Stringent conditions" refers to conditions that allow for the hybridization of substantially resided nucleic acids, where relatedness is a function of the sequence of nucleotides in the respective nucleic acids. For instance, for a nucleic acid of 100 nucleotides, such conditions will generally allow hybridization thereto of a second nucleic acid having at least about 85% homology, preferably having at least about 90% homology. Such hybridization conditions are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press. 1989.

The invention further relates to REST proteins and to proteins having sufficient zinc finger domains to confer the ability to bind the RE1 promoter element. Preferably, the protein has at least 4 of the zinc finger domains REST, more preferably at least 6, yet more preferably at least 7. Still

more preferably, the RE1 binding protein has all of the zinc finger domains. Preferably, the protein has the sequence of a contiguous stretch of at least about 252 amino acids of SEQ ID NO:1, more preferably, of a contiguous stretch of at least about 504 amino acids.

As discussed above, deletional or mutational methods of producing recombinant proteins that retain a given activity are well known. Thus, the embodiments of the present invention that relate to proteins also encompass analogs of REST proteins that retain one or both of the ability to bind the RE1 promoter element and to suppress the activity of a promoter to which the protein is bound. These analogs preferably lack no more than about 360 amino acid residues of deleted sequence at the C-terminal or N-terminal ends, more preferably no more than about 180 amino acid residues of deleted sequence. The remaining sequence of the REST protein will preferably have no more than about 20 point mutations, preferably no more than about 10 point mutations, more preferably no more than about 5 point mutations. The point mutations are preferably conservative point mutations. Preferably, the analogs will have at least about 85% homology, preferably at least about 90% homology, more preferably at least about 95% homology to a portion of an animal REST protein retaining one or both of REST's known activities, such as the proteins of SEQ ID NO:1 or SEQ ID NO:2.

Antigens for eliciting the production of antibodies against the REST protein can be produced recombinantly by expressing all of or a part of the nucleic acid of a REST protein in a bacteria or a yeast or other eukaryotic cell line. In one embodiemnt, the recombinant protein is expressed as a fusion protein, with the non-REST portion of the protein serving either to facilitate purification or to enhance the immunogenicity of the fusion protein. For instance, the non-REST portion comprises a protein for which there is a readily-available binding partner that is utilized for affinity purification of the fusion protein. The antigen includes an "antigenic determinant," i.e., a minimum segment of amino acids sufficient to bind specifically with an anti-REST antibody.

Rules for designing PCR primers are well known in the art, as reviewed by PCR Protocols, Cold Spring Harbor Press, 1991. Degenerate primers, i.e., preparations of primers that are heterogeneous at given sequence locations, are designed to amplify nucleic acid sequences that are highly related to, but not identical to, a REST protein. For instance, such degenerate primers, in one embodiment, are designed from the human REST cDNA and used to amplify nucleic acid sequences for REST proteins from non-human species, as illustrated in the examples.

The method by which human REST cDNA was isolated, which is described in detail in the examples, illustrates how readily RE1-binding domains from REST proteins are identified. In the isolation method, a library was made of cDNA from a REST-expressing cell and inserted into a

yeast expression vector for the GAL4 activation domain so that the library would express fusion proteins having one part derived from cDNA and another part that is the GAL4 activation domain. Initial partial cDNA clones were identified by their ability to bind an RE1 element on the promoters for two reporter genes and activate expression of those genes by causing the fused GAL4 activation domain to act on the promoters. These initial clones were of portions of the RE1 binding domain of the human REST protein. The same methodology can be used to identify other sequences from other animal sources that are sufficient to bind the RE1 element.

Additionally, the mutational and deletional methodologies that are well known in the art are applied to nucleic acids having the sequence of SEQ ID NO:2, which encodes the zinc finger domain of human REST. Nucleic acid constructs that express such mutated or deleted zinc finger domains are tested for the RE1 binding activity of the expressed protein. One facile method of doing this is to sub-clone the constructs into the GAL4 vector discussed above. Successful constructs activate the two RE1-containing reporter genes that were used in the initial cloning of human REST cDNA.

For identifying the suppressor domain of REST, one approach is to take a REST cDNA and create deletional mutants lacking segments at either the 5' or the 3' end by, for instance, partial digestion with S1 nuclease, Bal 31 or Mung Bean nuclease (the latter approach described in literature available from Stratagene, San Diego, CA, in connection with a commercial deletion cloning kit). Alternatively, the deletion mutants are constructed by subcloning restriction fragments of REST cDNA. The deletional constructs are cloned into expression vectors and tested for their ability to suppress the expression of a promoter that has a functional RE1 element. For instance, a reporter construct having the promoter for the CNS-type voltage-dependent sodium channel linked to the gene for chloramphenicol acetyl transferase ("CAT") is used. Such a vector is described below in the examples. Functional constructs diminish the level of expression of CAT, an enzyme that is readily measurable by well established techniques. See, for example, Gorman et al., Mol. Cell. Biol. 2, 1044-1051, 1982 and Young et al., DNA 4, 469-475, 1985.

Mutational and deletional approaches are applied to all of the nucleic acid sequences of the invention that express REST-related proteins. As discussed above, conservative mutations are preferred. Such conservative mutations include mutations that switch one amino acid for another wathin one of the following groups:

- 1. Small aliphatic, nonpolar or slightly polar residues: Ala, Ser, Thr. Pro and Gly;
- 2. Polar, negatively charged residues and their amides: Asp. Asn, Glu and Gln;
- 3. Polar, positively charged residues: His, Arg and Lys;
- 4. Large aliphatic, nonpolar residues: Met, Leu, Ile, Val and Cys; and

10

15

20

5. Aromatic residues: Phe, Tyr and Trp.

A preferred listing of conservative substitutions is the following:

Original Residue	Substitution
Ala	Gly, Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala, Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Tyr, Ile
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp _	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

The types of substitutions selected may be based on the analysis of the frequencies of amino
25 acid substitutions between homologous proteins of different species developed by Schulz et al.,

Principles of Protein Structure, Springer-Verlag, 1978, pp. 14-16, on the analyses of structureforming potentials developed by Chou and Fasman, Biochemistry 13, 211, 1974 or other such
methods reviewed by Schulz et al., Principles in Protein Structure, Springer-Verlag, 1978, pp.
108-130, and on the analysis of hydrophobicity patterns in proteins developed by Kyte and

30 Doolittle, J. Mol. Biol. 157: 105-132, 1982.

Numerous methods for determining percent homology are known in the art. One preferred method is to use version 6.0 of the GAP computer program for making sequence comparisons. The program is available from the University of Wisconsin Genetics Computer Group and utilizes the alignment method of Needleman and Wunsch, J. Mol. Biol. 48, 443, 5 1970, as revised by Smith and Waterman Adv. Appl. Math. 2, 482, 1981.

Nucleic acid molecules that bind to a REST-encoding nucleic acid under high stringency conditions are identified functionally, using methods outlined above, or by using the hybridization rules reviewed in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, 1989.

Antisera to REST are made by creating a REST antigen by linking a portion of the cDNA for human REST to a cDNA for glutathione s-transferase ("GST") found on a commercial vector. The resulting vector expresses a fusion protein containing an antigenic portion of REST and GST that is readily purified from the expressing bacteria using a glutathione affinity column. The purified antigenic fusion protein is used to immunize rabbits.

The same approach is used to make antigens based on other portions of the REST protein. Procedures for making antibodies and for identifying antigenic portions of proteins are well

known. See, for instance, Harlow, Antibodies, Cold Spring Harbor Press, 1989.

The proteins of the invention are made, in one embodiment, using the identical approach as for generating REST antisera. The cDNA specific for a given REST protein or analog thereof is linked using standard means to a cDNA for GST, found on a commercial vector, for example. The fusion protein expressed by such a vector construct includes the REST protein or analog and GST, and can be treated as above for purification. Should the GST segment of the fusion protein interfere with function, it is removed by partial proteolytic digestion approaches that preferentially attack unstructured regions, such as the linkers between GST and the REST-derived protein. The linkers are designed to lack structure, for instance using the rules for secondary structure-forming potential developed by Chou and Fasman, Biochemistry 13, 211, 1974. The linker is also designed to incorporate protease target amino acids, such as, for trypsin, arginine and lysine residues. To create the linkers, standard synthetic approaches for making oligonucleotides are employed together with standard subcloning methodologies. Other fusion partners other than GST can be used.

Also, of course, the REST proteins can be directly synthesized from nucleic acid (by the cellular machinery) without use of fusion partners. For instance, nucleic acids having the sequence of SEQ ID NO:10 are subcloned into an appropriate expression vector having an

appropriate promoter and expressed in an appropriate organism. (Note that REST lacks consensus glycosylation sites and, especially since it is not a membrane or exported protein. should lack glycosylations.) Antibodies against REST are employed to facilitate purification.

Additional purifications techniques are applied as needed, including without limitation, 5 preparative electrophoresis, FPLC (Pharmacia, Uppsala, Sweden), HPLC (e.g., using gel filtration, reverse-phase or mildly hydrophobic columns), gel filtration, differential precipitation (for instance, "salting out" precipitations), ion-exchange chromatography and affinity chromatography (including affinity chromatography using the RE1 duplex nucleotide sequence as the affinity ligand).

A protein or nucleic acid is "isolated" in accordance with the invention in that the molecular cloning of the nucleic acid of interest, for example, involves taking a human REST nucleic acid from a human cell, and isolating it from other human-derived nucleic acids. This isolated nucleic acid may then be inserted into a host cell, which may be yeast or bacteria, for example, or another human cell. A protein or nucleic acid is "substantially pure" in accordance 15 with the invention if it is predominantly free of other proteins or nucleic acids, respectively. A macromolecule, such as a nucleic acid or a protein, is predominantly free if it constitutes at least about 50% by weight of the given macromolecule in a composition. Preferably, the protein or nucleic acid of the present invention constitutes at least about 60% by weight of the total proteins or nucleic acids, respectively, that are present in a given composition thereof, 20 more preferably about 80%, still more preferably about 90%, yet more preferably about 95%, and most preferably about 100%. Such compositions are referred to herein as being proteins or nucleic acids that are 60% pure, 80% pure, 90% pure, 95% pure, or 100% pure, any of which are substantially pure.

One aspect of the present invention is directed to the use of "antisense" polynucleic 25 acid to treat neural diseases, including de-differentiated neural tumor cells and diseases characterized by diminished neural activity. Such an approach is also used to trigger the differentiation of neural stem cells. The approach involves the use of an antisense molecule designed to bind nascent mRNA (or "sense" strand) for a REST protein, thereby stopping or inhibiting the translation of the mRNA, or to bind to the REST gene to interfere with its 30 transcription. Once the sequence of the mRNA sought to be bound is known, an antisense molecule is designed that binds the sense strand by the Watson-Crick base-pairing rules. forming a duplex structure analogous to the DNA double helix. Gene Regulation: Biology of Antisense RNA and DNA, Erikson and Ixzant, eds., Raven Press, New York, 1991.

A serious barrier to fully exploiting this technology is the problem of efficiently introducing into cells a sufficient number of antisense molecules to effectively interfere with the translation of the targeted mRNA or the function of DNA. One method that has been employed to overcome this problem is to covalently modify the 5' or the 3' end of the antisense 5 polynucleic acid molecule with hydrophobic substituents. These modified nucleic acids generally gain access to the cells interior with greater efficiency. See, for example, Boutorin et al., FEBS Lett. 23,1382-1390, 1989; Shea et al, Nucleic Acids Res. 18, 3777-3783, 1990. Additionally, the phosphate backbone of the antisense molecules has been modified to remove the negative charge (see, for example, Agris et al., Biochemistry 25, 6268, 1986; Cazenave and 10 Helene in Antisense Nucleic Acids and Proteins: Fundamentals and Applications, Mol and Van der Krol, eds., p. 47 et seq., Marcel Dekker, New York, 1991) or the purine or pyrimidine bases have been modified (see, for example, Antisense Nucleic Acids and Proteins: Fundamentals and Applications, Mol and Van der Krol, eds., p. 47 et seq., Marcel Dekker, New York, 1991; Milligan et al. in Gene Therapy For Neoplastic Diseases, Huber and Laso, 15 eds., p. 228 et seq., New York Academy of Sciences, New York, 1994). Other attempts to overcome the cell penetration barrier include incorporating the antisense polynucleic acid sequence into an expression vector that is inserted into the cell in low copy number, but which, when in the cell, directs the cellular machinery to synthesize more substantial amounts of antisense polynucleic molecules. See, for example, Farhood et al., Ann. N.Y. Acad. Sci. 716, 20 23, 1994. This strategy includes the use of recombinant viruses that have an expression site into which the antisense sequence has been incorporated. See, e.g., Boris-Lawrie and Temin, Ann. N.Y. Acad. Sci., 716:59 (1994). Others have tried to increase membrane permeability by neutralizing the negative charges on antisense molecules or other nucleic acid molecules with polycations. See, e.g. Wu and Wu, Biochemistry, 27:887-892, 1988; Behr et al., Proc. Natl. Acad Sci U.S.A. 86:6982-6986, 1989.

The polynucleotide or nucleic acid compositions of the invention can be administered orally, topically, rectally, vaginally, by pulmonary route by use of an aerosol, or parenterally, i.e. intramuscularly, intraventricularly, subcutaneously, intraperitoneally or intravenously. The polynucleotide compositions are administered alone, or they are combined with a pharmaceutically-acceptable carrier or excipient according to standard pharmaceutical practice. For the oral mode of administration, the polynucleotide compositions are used in the form of tablets, capsules, lozenges, troches, powders, syrups, elixirs, aqueous solutions and

suspensions, and the like. In the case of tablets, carriers that are used include lactose, sodium citrate and salts of phosphoric acid. Various disintegrants such as starch, and lubricating agents such as magnesium stearate, sodium lauryl sulfate and talc, are commonly used in tablets. For oral administration in capsule form, useful diluents are lactose and high molecular weight 5 polyethylene glycols. When aqueous suspensions are required for oral use, the polynucleotide compositions are combined with emulsifying and suspending agents. If desired, certain sweetening and/or flavoring agents can be added. For parenteral administration, sterile solutions of the conjugate are usually prepared, and the pH of the solutions are suitably adjusted and buffered. For intravenous use, the total concentration of solutes is controlled to 10 render the preparation isotonic. For ocular administration, ointments or droppable liquids may be delivered by ocular delivery systems known to the art, such as applicators or eye droppers. Such compositions include mucomimetics, such as hyaluronic acid, chondroitin sulfate, hydroxypropyl methylcellulose or poly(vinyl alcohol), preservatives, such as sorbic acid or EDTA, and the usual quantities of diluents and/or carriers well known in the art. For 15 pulmonary administration, diluents and/or carriers are selected so as to allow the formation of an aerosol.

Generally, the polynucleotide compositions are administered in an effective amount.

An effective amount is an amount effective to either (1) reduce the symptoms of the disease sought to be treated or (2) induce a pharmacological change relevant to treating or preventing the disease sought to be treated.

For viral gene therapy vectors, dosages are generally from about 1 µg to about 1 mg of nucleic acid per kg of body mass. For non-infective gene therapy vectors, dosages are generally from about 1 µg to about 100 mg of nucleic acid per kg of body mass. Antisense oligonucleotide dosages are generally from about 1 µg to about 100 mg of nucleic acid per kg of body mass.

The invention also encompasses the use of gene therapy approaches to insert a gene expressing an RE1 binding domain but not a suppressor domain into de-differentiated tumor cells or neural cells with diminished neural activity. Gene therapy approaches for inserting a gene for a protein with REST activity into overactive neural cells are also within the invention.

30 Also, gene therapy approaches for inserting a gene for a REST suppressor domain linked to a promoter binding element to suppress the activity of the promoter bound by the binding element are also within the invention.

For gene therapy, medical workers prefer to incorporate, into one or more cell types of an organism, a DNA vector capable of directing the synthesis of a protein missing from the cell or useful to the cell or organism when expressed in greater amounts. The methods for introducing DNA to cause a cell to produce a new protein or a greater amount of a protein are called "transfection" methods. See, generally, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989.

A number of the above-discussed methods of enhancing cell penetration by antisense nucleic acid are generally applicable methods of incorporating a variety of nucleic acids into cells. Other general methods include calcium phosphate precipitation of nucleic acid and incubation with the target cells (Graham and Van der Eb, Virology, 52:456, 1983), coincubation of nucleic acid, DEAE-dextran and cells (Sompayrac and Danna, Proc. Natl. Acad. Sci., 12:7575, 1981), electroporation of cells in the presence of nucleic acid (Potter et al., Proc. Natl. Acad. Sci., 81:7161-7165, 1984), incorporating nucleic acid into virus coats to create transfection vehicles (Gitman et al., Proc. Natl. Acad. Sci. U.S.A., 82:7309-7313, 1985) and incubating cells with nucleic acid incorporated into liposomes (Wang and Huang, Proc. Natl. Acad. Sci., 84:7851-7855, 1987). An approach in employing gene therapy is to incorporate the gene sought to be introduced into the cell into a virus, such as an adenovirus. See, for instance, Akli et al., Nature Genetics 3, 224, 1993.

The stem cells that are useful in neural stem cell replacement therapy include human mesencephalic fetal brain cells, porcine fetal brain cells, human subventricular zone cells and glial progenitor cells, including O2A cells (which are progenitors for all glial cell types, including astrocytes and oligodendrocytes).

The invention also relates to methods of measuring a REST protein or mRNA from a tissue or staining a tissue for a REST protein or mRNA. Useful methods of measuring mRNA include Southern blot analysis, dot blot analysis, nuclear transcription analysis, histochemical staining for mRNA and polymerase chain reaction amplification methods. See generally, Ausubel et al., Current Protocols in Molecular Biology, Wiley Press, 1993; PCR Protocols, Cold Spring Harbor Press, 1991; and Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989. For in situ nucleic acid hybridization techniques, see Baldino et al., Methods in Enzymology 168, 761-777, 1989; Meson et al., Methods in Enzymology 168, 753-761, 1989; Harper et al., Methods in Enzymology 151, 539-551, 1987; Angerer et al., Methods in Enzymology 152, 649-661, 1987; Wilcox et al., Methods

30

compound(s).

in Enzymology 124, 510-533, 1986. Methods of measuring protein in a tissue include enzymelinked immunoassays ("ELISA"), immuno-diffusion assays, radio-immunoassays, immunoelectrophoresis, Western blot analyses and immunohistochemical staining techniques. See generally, Ausubel et al., Current Protocols in Molecular Biology, Wiley Press, 1993; 5 Antibodies, a Laboratory Manual, Cold Spring Harbor Press, 1988; and Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989.

PCR methods of amplifying nucleic acids utilize at least two primers. One of these primers is capable of hybridizing to a first strand of the nucleic acid to be amplified and of priming enzyme-driven nucleic acid synthesis in a first direction. The other is capable of 10 hybridizing the reciprocal sequence of the first strand (if the sequence to be amplified is single stranded, this sequence is initially hypothetical, but is synthesized in the first amplification cycle) and of priming nucleic acid synthesis from that strand in the direction opposite the first direction and towards the site of hybridization for the first primer. Conditions for conducting such amplifications, particularly under preferred high stringency conditions, are well known. 15 See, for example, PCR Protocols, Cold Spring Harbor Press, 1991.

The samples that are amenable to assaying or staining for REST protein or nucleic acid include, without limitation, cells or tissues (including nerve tissues), protein extracts, nucleic acid extracts and biological fluids such as cerebral fluid, serum and plasma. Preferred samples are nervous system-derived samples.

In screening assays for antagonists of the activity of REST, the agents to be screened include a great variety of chemicals including, but not limited to, biologically active molecules such as peptides, carbohydrates, alkaloids, aromatic compounds, polynucleotides and analogs thereof (particularly analogs that have been rendered more membrane permeable), DNA intercolating compounds and other pharmaceutical agents. One cell-free assay comprises the 25 steps of:

providing a nuclear extract. providing a REST protein, providing the nucleotide triphosphates necessary for transcription, providing a promoter sequence that includes an element effective to bind to REST and thereby be inhibited, providing a candidate compound or a cocktail of candidate compounds. mixing the extract, protein, promoter, nucleotide triphosphates, and candidate

incubating the mixture to allow transcription to proceed, and determining the level of the resulting transcription from the promoter, relative increases in transcription reflecting an inhibition of either the binding of REST to the promoter element or the activity of the suppressor domain of REST.

- For nuclear extracts from REST-expressing cells, the extract itself will generally provide sufficient amounts of the REST protein. Sufficient amounts of the nucleotide triphosphates may also be found in the nuclear extract; however, generally, additional nucleotide triphosphates are added to reduce the variability of the assay. The level of transcription is determined by primer extension as described by Bodner and Karin, Cell 50, 267-275, 1987.
- One embodiment of the cellular assay comprises the steps of:
 providing a eukaryotic cell line that expresses the REST protein (either natively or through a stable or transient transfection),
 providing a suitable medium for maintaining the cell line,
 adding to the medium a candidate compound or a cocktail of candidate compounds,
 incubating the cells to allow transcription to proceed, and
 determining the level of transcription from a REST-responsive promoter.

One way of determining the level of transcription is to have provided the cells with a REST-responsive promoter coupled to a gene for a readily measurable gene product. This method is, of course, indirect, since it requires the transcript, which one would prefer to directly measure, to be translated into a protein that is then measured. Nonetheless, the method is widely recognized as a surrogate measure of transcription. The appropriate RNA transcript is also measured by methods well known in the art, such as dot-blot hybridization or by Northern Blot analysis.

The REST protein has a negative influence on the activity of many promoters having an RE1 or an RE1-like sequence (such as that of the promoter for SCG10). Direct cloning strategies for such negative factors are difficult since they require time consuming measurements of the loss of a property. To create a positive signal that can more facilely be used to screen a cDNA library for REST-related cDNAs, a HeLa cell cDNA library was created to express fusion proteins between cDNA-encoded polypeptides and the activation domain of the yeast GAL4 regulatory protein. The library was designed to identify a clone encoding a fusion protein having an RE1-binding domain and a GAL4 activation domain. Such a fusion protein acts as a positive transcription factor on appropriate RE1-containing promoter.

WO 96/29433

A HeLa cell library was selected because HeLa cells do not express the type II voltage dependent sodium channel and express an RE1-binding activity.

The invention is described in more detail, but without limitation, by reference to the examples set forth below.

-5

Example 1 - "One-Hybrid" Cloning of Three Partial Sequences

a. Yeast Strains

The cloning strategy employed yeast containing two reporter genes having RE1 regulatory sequences in or adjacent to their promoters. One reporter gene was HIS3, which confers to yeast the ability to grow in media that lacks the amino acid histidine, functionally attached to the yeast GAL1 promoter. The GAL1 promoter is normally inactive in the absence of a yeast activator protein such as GAL4. The other reporter gene was the bacterial lac z gene functionally coupled to the yeast CYC1 promoter. The CYC1 promoter is normally inactive in the absence of a yeast activator protein such as GAL4.

15

i. The HIS3 Construct

Four copies of the 28 bp RE1 nucleic acid, SEQ ID NO:29, which had been synthesized by standard oligonucleotide synthesis methods, were cloned into a unique EcoRI site on yeast expression shuttle vector pTH1 (described by Flick and Johnson, Mol. Cell. Biol. 10(9), 4757-4769, 1990). The EcoRI site is adjacent (and 5') to a yeast GAL1 promoter that is functionally linked to a HIS3 gene. The shuttle vector also contained a marker gene that directed the expression of a gene that confers to yeast the ability to grow in the absence of the pyrimidine base uracil. A derivative plasmid containing four properly oriented copies of the RE1 sequence, as confirmed by sequence analysis, was isolated and designated pJAC12.

25

ii. The Lac z Construct

Four copies of the 28 bp RE1 nucleic acid, SEQ ID NO:29, were cloned between the Pst and BamHI sites upstream of the CYC1 promoter found on expression vector pCZi3gal (described by Lue and Kornberg, *Proc. Natl. Acad. Sci. USA* 84, 8839-8843, 1993), which promoter is functionally linked to a bacterial *lac z* gene. The vector also contained a marker gene that directed the expression of a gene that confers to yeast the ability to grow in the absence of the amino acid tryptophan. A derivative plasmid containing four properly oriented

copies of the RE1 nucleic acid, as confirmed by sequence analysis, was isolated and designated pJAC13.

iii. Yeast Transformation To Incorporate Reporter Genes

The reporter plasmids were linearized and introduced sequentially into a standard yeast strain (strain W303) by the LiAc method (Schiestl and Geitz, Curr. Gen. 16, 339-346, 1989). Transformants were selected by growth on plates lacking uracil (indicating the integration of pJAC12) and tryptophan (indicating the integration of pJAC13). Small scale preparations of total yeast genomic DNA were prepared from four colonies according to the method of Sherman et al., Methods in Yeast Genetics, Cold Spring Harbor Press, 1986, to confirm integration of the pJAC12 and pJAC13 reporter vectors into the yeast genome by Southern blot analysis using the RE1, CYC1 promoter, HIS3 gene, and TRP1 gene as probes. One of these four transformants was then utilized for the subsequent cDNA library transformation. This reporter strain was assessed for growth on his plates and screened for β-galactosidase activity and, as expected, was negative for both markers.

iv. Control Reporter Strain

By the same methods described above, a control strain derived from W303 was created that incorporated analogs of pJAC12 and pJAC13, wherein the RE1 nucleic acids were substituted with four copies of the inactive mutant RE1 nucleic acid, SEQ ID No. 30, described by Kraner et al., Neuron 9, 37-44, 1992.

b. cDNA Cloning

A HeLa cell cDNA library was constructed using the pGADGH plasmid containing the GAL4 activation domain (see Li and Herskowitz, Science 262: 1870-1874, 1993) functionally linked to a GAL4 promoter and having a polylinker site (including EcoRI and XhoI sites), located downstream of the activator domain sequence for inserting the cDNA. The library plasmid contains a marker for the ability to grow in the absence of the amino acid leucine. The library was linearized and introduced into the yeast reporter strain by the LiAc method. The cells were plated in leucine minus and histidine minus agar plates to select colonies that are putatively transformed with a cDNA to express an fusion protein having an RE1 binding domain (derived from cDNA) and a GAL4 activation domain.

One hundred his + colonies were impressed onto filter paper and permeabilized by freeze-thawing. The filter paper was layered onto another filter paper containing the β galactosidase substrate 5-bromo-4-chloro-3-indoyl-b-D-galactoside (X-gal, available from Sigma Chemical Co., St. Louis). The filter paper was incubated at room temperature and monitored 5 for blue spots, which indicate β -galactosidase positive colonies. Four colonies that were positive for the lac z marker were isolated. Plasmids containing the cDNA from these four colonies was isolated as described by Bartel et al., in Cellular Interactions in Development: A Practical Approach, D.A. Hartley, ed., New York: Oxford University Press, 1994, pp 53-179, and amplified in bacteria. The plasmids were introduced into the control yeast strain (wherein 10 the reporter gene promoters contained mutant RE1 sequences). Three of the four plasmids failed to transform the control strain, indicating that the fusion proteins they encoded interacted specifically with the RE1 nucleic acid. These plasmids were designated p73, p90 and p613. The three insert cDNAs were sequenced by the chain termination method (Sanger et al., Proc. Natl. Acad. Sci. USA 74, 998-1002, 1977) and found to include the sequences of SEQ ID 15 NO:3, SEQ ID NO:4 and SEQ ID NO:5, all of which encode overlapping portions of an apparent zinc-finger DNA-binding domain (nucleotides 216-1622, 636-1725 and 695-1622 of Fig. 1, respectively).

Example 2 - Cloning of Two Overlapping Sequences Encoding REST

SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5 were used to probe another HeLa cell cDNA library that was cloned into the Lambda Zap II phage (Stratagene, Inc., San Diego, CA). Two phage isolates containing overlapping cDNAs of 3082 and 4408 bp were isolated (phages NH2 and NH7, respectively). These cDNAs are designated SEQ ID NO: 6 and SEQ ID NO:7 and encode nucleotides -175-1616 and 1472-5324 of Fig. 1, respectively. From the overlap of these two cDNAs, most of the full length REST cDNA can be deduced. The 5' segment, up to position -325, was determined by applying the 5' RACE PCR technique to HeLa cell cDNA. This segment is designated SEQ ID NO:1. The deduced amino acid sequence of REST is shown is Figure 1. Note that Lambda Zap II is readily convertible to the Bluescript plasmid using EcoRI as outlined by the supplier.

Example 3 - Expression of REST Antigen and Polyclonal Antibody Production

For example 3, a 1.5 kilobase EcoRI-XhoI fragment of p73 comprising all of SEQ ID NO:3 was cloned in phase with the cDNA for glutathione s-transferase ("GST") in the

commercial vector pGEX4T3 (Pharmacia, Uppsala, Sweden). The GST-REST fusion protein was produced in E.coli strain XL-1 blue (Stratagene, San Diego, CA) and purified on a glutathione-Sepharose column (Pharmacia, Uppsala, Sweden). The purified fusion protein was used to immunize two rabbits (Pocono Rabbit Farms, PA) to produce a polyclonal antibody preparation against REST.

Example 4 - RNA Hybridization (Northern Blots)

Total cellular RNA from HeLa cells, PC12 cells, L6 skeletal muscle cells and dorsal root ganglion was isolated as described by Toledo-Aral et al., Neuron, in press) and 10 poly-A+-selected using a commercially available kit (Pharmacia, Inc., Uppsala, Sweden). Messenger RNA (2-4 μ g) was fractionated on denaturing gels and then electrophoretically transferred onto nylon paper for hybridization. A DNA probe of human REST was generated by random primer labeling of the EcoRI - XhoI fragment of p73, which includes the nucleic acid of SEQ ID NO:3, to incorporate ³²P. A rat REST cDNA (600 bp) was obtained by PCR 15 (with an initial reverse-transcriptase step) of rat skeletal muscle mRNA using a degenerate primer modelled on the sequence of amino acids 146 to 153 (nucleotides 481 to 504) of the plus strand of SEQ ID NO:1 and a degenerate primer modelled on the amino-acid-encoding sequence of amino acid residues 363 to 370 (nucleotides 1087 to 1110) of the minus strand of SEQ ID NO:1. The PCR-amplified cDNA was cloned into pGEM-7Z (Promega, Madison, 20 WI), and workable amounts of the plasmid were grown in bacteria. A rat REST riboprobe was manufactured by linearizing the plasmid with Accl and transcribing it with T7 polymerase in the presence of ³²P-UTP (Dupont, Wilmington, DE). A riboprobe for the CNS-type sodium channel was made as described by D'Arcangelo et al., J. Cell Biol., 10(9), 4757-4769, 1993. Hybridization and washing conditions used with the rat REST and sodium channel riboprobes were as described by Toledo-Aral et al., Neuron, in press; for the human REST DNA probe. the hybridization and washing solutions were the same as those used for the riboprobes, except that the blots were hybridized at 37°C and washed at 32°C.

Northern blot analysis for mRNS for the CNS-type sodium channel and REST in a number of cell types and tissues produced the following results:

10

Cell or Tissue Type	CNS-type Sodium Channel mRNA	REST mRNA
HeLa cells	none	high levels
rat L6 skeletal muscle cells	none	high levels
rat PC12 cells	high level	extremely low levels
mouse dorsal root ganglia	extremely low levels	high levels

Example 5 - Western Blot Analysis

Western immunoblots of proteins derived from nuclear extracts were performed according to standard procedures, as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harber Lab., Cold Spring Harbor, NY, 1989. Nuclear extracts were prepared by the single lysis method (Sambrook et al., 1989). Extracts were combined with an equal volume of 2X Laemmli sample buffer (Laemmli, Nature, 227, 680-15 685, 1970) and boiled for 15 minutes. Samples were resolved by SDS-PAGE on 7.5% gels, transferred to nitrocellulose, and the nitrocellulose was blocked with 10% milk in TTBS (Sambrook et al., 1989). Immunoblotting was performed using the enhanced chemiluminescence method using a commercial kit (Amersham, Burlington, MA). The antibody to REST-GST was used at a 1:20 dilution after purification by FPLC on an alkyl 20 Superose (a highly crosslinked agarose substituted with octyl groups) column (Pharmacia, Uppsala, Sweden).

Nuclear extracts were made from the PC12 cell line derived from a neural pheochromocytoma, which expresses the CNS-type voltage-dependent sodium channel and does not express an RE1 binding activity, and from HeLa cells, which do not express the CNS-type 25 voltage-dependent sodium channel and do express an RE1 binding activity. Western blots probed with the polyclonal antibodies to human REST indicated the presence of an immunoreactive protein of molecular weight 121 kDa in HeLa cell nuclear extracts, but no immunoreactive protein in PC12 cell nuclear extracts.

30 Example 6 - In Situ Hybridization

The developmental pattern of expression of REST was analyzed by in situ hybridization in mouse embryos. A 600 bp fragment of mouse REST cDNA (encompassing most of the zinc finger domain) was prepared from 8.5 day mouse embryos by the PCR method described in

Example 4 for the preparation of rat REST cDNA. The amplification product was cloned into a Bluescript vector (Stratagene, San Diego, CA) and partially sequenced using the Sequenase Kit (US Biochemicals, Cleveland, OH). In situ hybridization of intact embryos using digoxigenin (DIG-11-UTP, available from Boehringer Mannheim) labeled RNA probes for 5 mouse Hox-B1 (Frohman at al., Development, 110, 589-608, 1990), and Gbx-2 (Frohman et al., Mouse Genome, 91, 323-325, 1993). Hybridization was performed using a published protocol (see Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, 1989). In brief, embryos were fixed overnight in paraformaldehyde, incubated in hydrogen peroxide to inactivate endogenous phosphatases, lightly proteinase K digested, 10 refixed, and hybridized at 70 °C in 1 ml of 50% formamide, 5 x SSC pH 4.5, 50 μ g/ml yeast RNA, 1% SDS, 50 μ g/ml heparin, 0.1% CHAPS, and 5mM EDTA containing 1 μ g of probe. The embryos were rinsed in a low wash solution (50% formamide, 5 x SSC, pH 4.5, 1% SDS, 0.1% CHAPS; 70°C), treated with RNAse A, rinsed with a high stringency wash solution (50% formamide, 2 x SSC, pH 4.5, 0.1% CHAPS; 65°C), and incubated with an 15 alkaline-phosphatase coupled rabbit anti-digoxin antisera (Boehringer Mannheim, Indianapolis, IN) The enzyme activity of the reporter was detected by a color reaction with 5-bromo-4chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT), which resulted in the deposition of a water-insoluble purple precipitate. Embryos were rinsed, washed into 80% glycerol, and photographed intact and in slices.

The *in situ* hybridization results for 9.5 day embryos indicated the presence of abundant REST mRNA in all tissues except the developing brain and spinal cord. Robust expression of REST mRNA was found in neural crest-derived dorsal root ganglia, indicating the expression of REST in some non-CNS neural tissue.

25 Example 7 - Mobility Shift Assays for Proteins That Bind RE1 Sequences

The presence of RE1 binding activity in various cells and tissues was tested using a gel mobility shift assay. Nuclear extracts from HeLa, L6, and primary cultures of rat embryonic skeletal muscle cells were prepared as described by Dignam et al., Nucl. Acids Res., 11, 1475-1489, 1983. The extracts were preincubated 15 minutes at room temperature with either buffer control, competitor DNA, REST-GST polyclonal antisera, or rabbit preimmune serum, and then incubated for two hours at room temperature with a 114 bp ³²P end-labeled DNA probe containing nucleotides -1051 to 837 of the 5' flanking sequence for the CNS-type sodium channel gene, which promotes sequence includes the RE1 sequence. The samples were

resolved by electrophoresis on a 5% non-denaturing polyacrylamide gel, which was then autoradiographed. The presence of binding was indicated by the presence of a DNA complex that moved more slowly in the gel than does the free DNA probe.

The results were that HeLa, L6 and rat embryonic skeletal muscle all contained an RE1 binding activity that was competed away with excess unlabelled RE1 containing DNA but not by DNA containing the inactive RE1 mutant described by Kraner et al., Neuron, 9, 37-44, 1992. The polyclonal antisera to the REST-GST fusion further retarded mobility, while pre-immune serum had no effect. This result indicates that a REST-like protein is responsible for the binding indicated by the gel shift assay.

10

Example 8 - Expression Vector Encoding The Complete Human REST Protein

The NH2 vector containing the nucleic acid of SEQ ID NO: 6 was digested with Hind III and Hinc II; and the NH7 vector containing the nucleic acid of SEQ ID NO:7 was digested with Hinc II and Bgl II. The excised inserts were subcloned into a Hind III and Barn HI digested pCMV I-amp (Invitrogen, Inc., San Diego) vector. The Hinc II digestion cleaved the overlap region of NH2 and NH7 at nucleotide 1575, allowing for a contiguous insert of nucleotides -175 through 3656 to be isolated.

Example 9 - Transfection Studies of REST Function

Transfert transfection of PC12 cells with a plasmid containing the chloramphenicol acetyl transferase (CAT) gene attached to the RE1-containing promoter for the CNS-type sodium channel results in the expression of CAT (the plasmid designated herein as "type II-CAT"). This plasmid has been described by Kraner et al., Neuron, 9, 37-44, 1992. A control CAT vector driven by the strong rous sarcoma virus (RSV) promoter has been described by Kraner et al., 1992 and Gorman et al., Proc. Natl. Acad. Sci. USA 79, 6777-6781, 1982. To test whether this expression could be shut-down by the REST protein, cotransfection experiments using the type II-CAT plasmid and a plasmid containing the REST cDNA coupled to the cytomegalovirus ("CMV") promoter were undertaken. A fragment of the REST cDNA, encoding the entire REST protein, with HindIII and BglI termini (including nucleotides -175 to 3656 of SEQ ID NO:1) was subcloned downstream of the CMV promoter in the commercial mammalian expression vector pCDNA l-amp (InVitrogen, Inc., San Diego, CA) between the HindIII and BamHI sites to create the CMV-REST vector. The resulting expression vector was designated REST-Express. Rat PC12 cells were transfected with 30 μg of REST-Express and

30 μg of either type II-CAT or RSV-CAT by electroporation (Kraner et al., 1992). Forty-eight hours after transfection the cells were harvested, centrifuged and lysed by freeze-thaw cycles. The supernatant was analyzed for CAT activity as previously described in Maue et al., Neuron, 4, 223-231, 1990. A cDNA encoding the Zn finger region of REST (including nucleotides 481 to 1236 of SEQ ID NO:1) was cloned independently into the pCDNA1-amp vector and was used as an interfering form of REST in transient transfection assays. L6 muscle cells and PC12 cells were transfected with 30 μg of the interfering REST vector along with 30 μg of type II-CAT plasmid by electroporation and treated as above.

The results were that co-transfection into PC12 cells of REST-Express along with the

type II-CAT resulted in a ten-fold decrease in activity versus the activity seen with type II-CAT

alone. REST-Express had not effect on the expression of CAT by RSV-CAT. The interfering

REST vector, encoding just the DNA binding domain of REST, had no effect on the expression

of type II-CAT in PC12 cells. However, in L6 muscle cells, which contain an endogenous

REST activity, the interfering REST vector derepressed the expression of type II-CAT, which

is otherwise inactive in L6 cells. This latter result is consistent with REST having a suppressor

function that is held in the vicinity of the promoter for the CNS-type sodium channel by the

DNA-binding domain. By competing the complete REST protein from the promoter, the

interfering form of REST — containing only the DNA-binding domain — de-represses the

promoter.

20

Example 10 - Localization of the Repressor Function

A number of restriction fragments were isolated from the full length expression clone described in Example 8 or from the NH2 clone and subcloned into the CMV-promoted expression vector also described in Example 8. Two other REST fragments were available from cDNA library screenings. These were clones NH10 and NH12, which contain nucleotides 121-1581 and 25-1308 of Figure 1, respectively (which sequences are designated SEQ ID NO:27 and 28). The inserts of these clones were excised with EcoRI and subcloned into the CMV-promoted vector. In total, the inserts subcloned into the expression vector had the following sequence from Figure 1:

30

- 1. Nucleotides 31-3976
- 2. Nucleotides 31-2234
- 3. Nucleotides 31-1940
- 4. Nucleotides 121-1581
- 5. Nucleotides 25-1308

6. Nucleotides 31-2491 and 2683-3976

In the last of these clones, the sequence between two BstXI restriction sites is excised. These subclones are co-transfected with PC12 cells along with the type II-CAT plasmid as described above to determine the silencing potential of the expressed fragment.

Example 11 - Designing PCR Amplification Primers

The PCR primers used to amplify sequences encoding amino acid residues 146 through 370 in Example 4 were designed as follows. First, the 146 to 153 sequence was translated into the following sequence-encoding nucleic acid sequence (SEQ ID NO:8):

TGYAARCCNTGYCARTAYGARGCN.

where Y = T/C, R = A/G and N = A/G/T/C. Next, the sequence of amino acid residues 363 to 370 was translated as above. This translated sequence was used to define the following opposite strand sequence (SEQ ID NO:9):

15 NGTYTTRTARTCRCARTGNGGRCA.

While this invention has been described with an emphasis upon preferred embodiments, it will be obvious to those of ordinary skill in the art that variations in the preferred compositions and methods may be used and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the claims that follow the Sequence Listing.

- 29 -

SEQUENCE LISTING

- (1) GENERAL INFORMATION
- (i) APPLICANT: Mandel, Gail, Chong, Jayhong A.
- 5 (ii) TITLE OF INVENTION: REST Protein and DNA
 - (iii) NUMBER OF SEQUENCES: 29
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Dechert Price & Rhoads
 - (B) STREET: P.O. Box 5218
- 10 (C) CITY: Princeton
 - (D) STATE: New Jersey
 - (E) COUNTRY: USA
 - (F) ZIP: 08543-5218
- (v) COMPUTER READABLE FORM:
- 15 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 - (B) COMPUTER: IBM-compatible
 - (C) OPERATING SYSTEM: DOS 5.0
 - (D) SOFTWARE: WordPerfect
 - (vi) CURRENT APPLICATION DATA:
- 20 (A) APPLICATION NUMBER:
 - (B) FILING DATE: March 23, 1995
 - (C) CLASSIFICATION:
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Allen Bloom
- 25 (B) REGISTRATION NUMBER: 29,135
 - (C) REFERENCE/DOCKET NUMBER: 317743-101 WO
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (609) 520-3214
 - (B) TELEFAX: (609) 520-3259
- 30 (2) INFORMATION FOR SEQ ID NO: 1:
 - (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 5648 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
- 35 (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
- 40 (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:

0.0000177	TO COTTO O CIO DO
O`96/29433	PCT/US96/039

- 30 -

	(A) LIBRARY: cDNA	
	(x) PUBLICATION INFORMATION:	
	(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José,	
	Toledo-Aral,	
5	Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yeler	ıa
	M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail	
	(B) TITLE: REST: A Mammalian Silencer Protein that Restric	ts
	Sodium Channel Gene Expression to Neurons	
	(C) JOURNAL: Cell	
10	(D) VOLUME: 80	
•	(E) ISSUE:	
	(F) PAGES:	
	(G) DATE: March 24, 1995	
	(K) RELEVANT RESIDUES IN SEQ ID NO:1:FROM -1 TO 5648	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
	ATCTGGCGCG GCGTAGCCCT GTGTTGGAAT GTGCGGCTGC CGCGAGCTCG	50
	CGGCGCAGCA GCGGAGCGAG CGCCGGGGGCC CCAGACCCTG	100
20		
	GCGGCGGCTG CGGCAGCCGA GACGGCAGGG CGAGGCCCGG AGGCCTGAGC	150
	ACCCTCTGCA GCCCCACTCC TGGGCCTTCT TGGTCCACGA CGGCCCCAGC	200
~~		
25	ACCCAACTTT ACCACCCTCC CCCACCTCTC CCCCGAAACT CCAGCAACAA	25
	AGAAAAGTAG TCGGAGAAGG AGCGGCGACT CAGGGTCGCC CGCCCCTCCT	30
	G100010011 0000011710 1077	
20	CACCGAGGAA GGCCGAATAC AGTT	324
30	NEG 000 NG 0NG 0EN NEG 000 0NG EGT EGT 000 000 000 000 0EG	~ ~ ~
	ATG GCC ACC CAG GTA ATG GGG CAG TCT TCT GGA GGA GGA GGG CTG	36
	Met Ala Thr Gln Val Met Gly Gln Ser Ser Gly Gly Gly Leu	
	1 5 10 15	
35	TTT ACC AGC AGT GGC AAC ATT GGA ATG GCC CTG CCT AAC GAC ATG	43.
"		414
	Phe Thr Ser Ser Gly Asn Ile Gly Met Ala Leu Pro Asn Asp Met	•
	20 25 30	
	TAT GAC TTG CAT GAC CTT TCC AAA GCT GAA CTG GCC GCA CCT CAG	45
40		* 7
+∪	Tyr Asp Leu His Asp Leu Ser Lys Ala Glu Leu Ala Ala Pro Gln	
	35 40 45	

	C THE															
	•	. All	ATC	CTC	GCA	LAA-1	GTG	GCC	C TT	A ACT	r GGG	GAA	GT	A AA	r GGC	50
	Let	1 TTE	'Met	Let	ı Ala	Asr	ı Val	Ala	a Let	Thi	Gly	Glu	[Va]	Ası	ı Gly	•
					50					55					60	
5	AGC	TGC	ייטיניי	י כי מי	· • • • • • • • • • • • • • • • • • • •	- cmc										
_		Cre	· 101	. DAI		CIG	GIC	GGT	GAZ	A GAZ	A AGA	CAG	ATO	GCZ	GAA	54
	001	Cyb	Cys	ASL		.Leu	val	GIŞ	/ Gli	ı Glu	Arg	Gln	Met	Ala	Glu	
					· 6 5					70					75	
	CTG	ATG	CCG	GTT	GGG	GAT	AAC	י אאר	لململ	י יייריא	ር እጥ	» Cm	~~~		GGA	
10	Leu	Met	Pro	Val	Gly	Asp	Asn	Asn	Phe	Ser	. Den	Sor Sor	Chi	GAA	GGA	594
					80					85	wah	261	GIU	GIU	•	
										03					90	
	GAA	GGA	CTT	GAA	GAG	TCT	GCT	GAT	ATA	. AAA	GGT	GAA	CCT	CAT	GGA	639
	Glu	Gly	Leu	Glu	Glu	Ser	Ala	Asp	Ile	Lys	Gly	Glu	Pro	His	Gly	
15					95					100	_				105	
	CTG	GAA	AAC	ATG	GAA	CTG	AGA	AGT	TTG	GAA	CTC	AGC	GTC	GTA	GAA	684
	Leu	Glu	Asn	Met	Glu	Leu	Arg	Ser	Leu	Glu	Leu	Ser	Val	Val	Glu	
					110					115					120	
20																
	CCT	CAG	CCT	GTA	TTT	GAG	GCA	TCA	GGT	GCT	CCA	GAT	ATT	TAC	AGT	729
	Pro	Gln	Pro	Val	Phe	Glu	Ala	Ser	Gly	Ala	Pro	Asp	Ile	Tyr	Ser	
					125					130					135	
25																
25	TCA	TAA	AAA	GCT-	CTT	GCC	CCT	GAA	ACA	CCT	GGA	GCG	GAG	GAC	AAA	774
	Ser	Asn	Lys	Ala	Leu	Ala	Pro	Glu	Thr	Pro	Gly	Ala	Glu	Asp	Lys	
					140					145					150	
20	GGC	AAG	AGC	TCG	AAG	ACC	AAA	CCC	TTT	CGC	TGT	AAG	CCA	TGC	CAA	819
30	GIY	Lys	Ser	Ser		Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	
					155					160					165	
							•									
	TAT	GAA	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC	ATC	AGA	GTT	864
25	ıyr	GIU	Ala	Glu		Glu	Glu	Gln	Phe	Val	His	His	Ile	Arg	Val	
35					170					175					180	
	C1 C														•	
				AAG												909
	nıs	ser	Ala	rys		Phe	Phe	Val	Glu	Glu	Ser	Ala	Glu	Lys	Gln	
40					185					190					195	
TU																

	GCA	AAA	GCC	AGG	GAA	TCT	GGC	TCT	TCC	ACT	GCA	GAA	GAG	.GGA	.GAT	954
	Ala	Lys	Ala	Arg	.Glu	Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
					.200					205					210	
5	TTC	TCC	AAG	GGC	CCC	ATT	CGC	TGT	GAC	CGC	TGC	GGC	TAC	AAT	ACT	999
	. Phe	Ser	Lys	Gly	Pro	Ile	Arg	Cys	Asp	Arg	Cys	Gly	Tyr	Asn	Thr	
					215					220					225	
	AAT	CGA	TAT	GAT	CAC	TAT	ACA	GCA	CAC	CTG	AAA	CAC	CAC	ACC	AGA	1044
10	Asn	Arg	Tyr	Asp	His	Tyr	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	
					230					235			-		240	
																1089
	Ala	Gly	Asp	Asn		Arg	Val	Tyr	Lys	Cys	Ile	Ile	Cys	Thr	Tyr	
15					245					250					255	
																1134
	Tnr	Thr	Val	Ser		Tyr	His	Trp	Arg		His	Leu	Arg	Asn		
20					260					265					270	
20	بلمثميات	CCD	NCC.	222	CTLX	ma.c		mam.	223							
																1179
	FIIC	FIO	Αij	Lys	275	ıyı	1111	Cys	GIY		Cys	ASI	Tyr	Pne		
					213					280					285	
25	GAC	AGA	מממ	אמר	מממ	ידמיני	CTT	CAG	Слт	CTTT	אכא	አርጥ	CNT	N C N	CCN	1224
											Arg					1224
		3	-7-		290	- , -	,	0211		295	9	****		****	300	
															300	
	GAA	CGC	CCA	TAT	AAA	TGT	GAA	CTT	TGT	ССТ	TAC	TCA	AGT	TCT	CAG	1269
30											Tyr					
				-	305	-					•				315	
	AAG	ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	ACT	CAT	TCA	GGT	GAG	AAG	1314
											His					
35					320					325			_		330	
	CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	1359
	Pro	Phe	Lys;	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
			•		335					340					345	
40																

	GAA.	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTI	CAC	AAT	GGG	CCI	AAA	CCT	1.404
	Glu	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	.Asr	Gly	Pro	Lys	Pro	
					350					355				-	360	
-5	لسلس	ייבב.	Дес	CCN	·C N C	·m~~										
																1449
	.204	'AGII	Cy.S	FIU	365		Asp	Tyr	ьys			Asp	Arg	Ser	Asn	
					363					370					375	
	TTC	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1494
10	Phe	Lys	Lys	His	Val	Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
					380					385				•	390	
	TGC	CCT	GTA	TGT	GAC	TAT	GCA	GCT	ፐርር	AAG	አአር	⊕C#	3 3 m	Cm»	0.0	1539
	Cys	Pro	Val	Cys	Asp	Tyr	Ala	Ala	Ser	Lve	Lve	Cve	YAT	LIA	CAG	1539
15		•		•	395	•				400	Dy S	cys	ASII	Leu	405	
															403	
	TAT	CAC	TTC	AAA	TCT	AAG	CAT	CCT	ACT	TGT	CCT	AAT	444	ACA	ATC	1584
	Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Cys	Pro	Asn	Lvs	Thr	Met	1304
					410					415			-,-		420	
20																
	GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	GCT	1629
•	Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	
					425					430					435	
25	GAC	TTG	CCT	GAT	AAT	ATT	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	1674
	Asp	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
					440					445					450	
20	ACA	AAA	ATA	AAA	GGG	GAT	GTG	GCT	GGA	AAG	AAA	TAA	GAA	AAG	TCC	1719
30	Thr	Lys	Ile	Lys	Gly	yab	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
					455					460					465	
	GTC	מממ	GC _A	GAG	מממ	אכא		CTC	man.							
								Val							TCT	1764
35		-,,,			470	9	rsp	Val	SeI		GIU	Lys	Lys	Pro		
					4,0					475					480	
	AAT	TAA	GTG	TCA	GTG	ATC	CAG	GTG	ACT	ACC	AGA	ACT	CGA	ααα	TCA	1800
								Val								2009
					485					490			3		495	
10																

	GTA	ACA	GAG	GTG	AAA	GAG	ATG	GAT	GTG	CAT	ACA	GGA	AGC	AAT	TCA	1854
											Thr					
					500					505					510	
`5																1899
	Glu	Lys	Phe	Ser	Lys	Thr	Lys	.Lys	Ser	Lys	.Arg	Lys	Leu	Glu	Val	
					515					520					525	
																1944
10	Asp	Ser	His	Ser	Leu	His	Gly	Pro	Val	Asn	Asp	Glu	Glu	Ser	Ser	
					530					535			•		540	
																1989
	Thr	Lys	Lys	Lys		Lys	Val	Glu	Ser	Lys	Ser	Lys	Asn	Asn	Ser	
15					545					550					555	
																2034
	Gin	GIu	Val	Pro		Gly	Asp	Ser	Lys	Val	Glu	Glu	Asn	Lys	Lys	
20					560					565					570	
20	C	מת ת	» Om	maa	1 ma											
																2079
	GIII	ASII	Int	Cys		Lys	Lys	Ser	Thr		Lys	Lys	Thr	Leu		
					575					580					585	
25	ידממ	מממ	TCA	λCT	אאכ	***	200	N C/M								
											Pro					2124
		273	501	Jer	590	Lys	SET	Ser	Lys		Pro	Gin	Lys	GIU		
					330					595					600	
	GTT	GAG	AAG	GGA	тст	GCT	CAG	ATC.	GAC	رحس		CNG	n TC	ccc	CCT	2169
30											Pro					2163
			-3-	2	605				p	610	110	GIII	Mec	Gly	615	
										010					013	
	GCT	CCC	ACA	GAG	GCG	GTT	CAG	AAG	GGG	CCC	GTT	CAG	GTG	GAG	СТС	2214
											Val					2217
35					620			-,-		625					630	
	CCA	CCT	ccc	ATG	GAG	CAT	GCT	CAG	ATG	GAG	GGT	GCC	CAG	ATA	CGG	2259
											Gly					
					635					640	•			_	645	
1 0																

	CCI	GC'	TCC	T GA	C GA	G CC.	r GT	CA(G AT	G GA	G GT	G GTT	ר רא	c cn	c cc	G :2304
	Pro	Al	a Pro	o As	p Gl	u Pro	Va:	l Gli	n 'Me	t Gl	u Va	l Vaï	ים. ומוי	n Cl		G .2304
					65	0				-65			- G1.	n GI		
											-				-66	U
:5	CCT	GC.	CAC	G AA	G GA	CTO	CTO	G CC1	ר ככי	T GT	c ca	ב ככי		T 63	~ ~~	G :2349
	Pro	Ala	a Glr	n Lys	s Gli	ı Let	ı Leı	ı Pro	Pro	. Va	l Gly	u Pro	. Bl.	CA	G AT	.2349
					665					670		ı PIC) AL	a G1:		
										0,1	J				67	5
	GTG	GG1	GCC	CAZ	ATT	GTA	CTI	, GC _M	י ראכ	י אידיני		·				2394
10	Val	Gly	/ Ala	Glr	ı Ile	Val	Leu	ıΔla	. Wic	· Mai	. C1.	Lev		r cc.	r ccc	2394
		_			680)		, AIG	nis			ı Lev	Pro	Pro	Pro	
										685	•				690)
	ATG	GAG	ACT	GCI	CAG	ACG	GAG	ייים:		. ~~						2439
	Met	Glu	Thr	Ala	Gln	Thr	Glu	Val	אוה	- CAA	ATG	: Gly	CCI	. GC.	ccc	2439
15					695		014	Val	MIG			GIY	Pro	Ala	Pro)
										700	,				705	;
	ATG	GAA	CCT	GCT	CAG	ATG	GAG	ىلىدات	CCC	C > C						2484
	Met	Glu	Pro	Ala	Gln	Met	Glu	Val	83.	CAG	GTA	GAA Glu	TCT	GCI	, ccc	2484
					710		GIU	VAI	ATA			Glu	Ser	Ala	Pro	ı
20					0					715					720	
	ATG	CAG	GTG	GTC	CAG	AAG	GAG	CCT	COVE	G) G						2529
	Met	Gln	Val	Val	Gln	Lvs	Glu	Dro	W-1	CAG	ATG	GAG Glu	CTG	TCT	CCT	2529
					725	2,5	GIU	PIO	vai		met	GIU	Leu	Ser	Pro	
										730					735	
25	CCC	ATG	GAG	GTG	GTC	CAG	AAG	GNC	CCm	com	~~					2574
	Pro	Met	Glu	Val	Val	Gln	Tare	Glu	D~c	GII	CAG	Ile	GAG	CTG	TCT	2574
				_	740		בינם	GIU	PIO		GIN	TIE	GIu	Leu		
					,					745					750	
	CCT	ccc	ATG	GAG	GTG	GTC	CAG	A A C	CNN	COM	C TOTAL					2619
30	Pro	Pro	Met	Glu	Val	Val	Gln	Lve	Glu	Dro	GII	Lys	ATA	GAG	CTG	2619
					755			273	Gru		vaı	гÀг	TIE	GIu		
										760					765	
	TCT	CCT	ccc	ATA	GAG	GTG	GTC	CAG	מממ	CNC			 -			2664
	Ser	Pro	Pro	Ile	Glu	Val	Val	Clo	Tue	Clu	CCI	Val	CAG	ATG	GAG	2664
35					770			G111	пуs		PIO	vai	Gin	Met		
					•					775					780	
	TTG '	TCT	CCT	CCC	ATC	GGG	CTC	مستات	C		~ ~					
	Leu !	Ser	Pro	Pro	Met	ദാഹ	773 10	27 J	CAG	AAG	GAG	CCT	GCT	CAG	AGG	2709
	Leu :		0	4 .	785	GT Å	val	AGT			GIu	Pro .	Ala	Gln	Arg	
					,00					790					795	

	GAG	CCA	CCT	CCT	CCC	,AGA	GAG	CCT	CCC	CTT	CAC	ATG	GAG	CCA	ATT	.2754
	Glu	Pro	Pro	Pro	Pro	Arg	.Glu	Pro	Pro	Leu	His	'Met	Glu	Pro	Ile	
					.800					805					·B10	
5	TCC	AAA	AAG	CCT	CCT	CTC	CGA	AAA	GAT	.AAA	AAG	GAA	AAG	TCT	AAC	2799
	Ser	Lys	Lys	'Pro	Pro	Leu	Arg	Lys	Asp	Lys	Lys	Glu	Lys	Ser	Asn	
					815					820					825	
	ATG	CAG	AGT	GAA	AGG	GCA	CGG	AAG	GAG	CAA	GTC	CTT	ATT	GAA	GTT	2844
10	Met	Gln	Ser	Glu	Arg	Ala	Arg	Lys	Glu	Gln	Val	Leu	Ile	Glu	Val	
					830		•			835			٠.		840	
	GGC	TTA	GTG	CCT	GTT	AAA	GAT	AGC	TGG	CTT	CTA	AAG	GAA	AGT	GTA	2885
	Gly	Leu	Val	Pro	Val	Lys	Asp	Ser	Trp	Leu	Leu	Lys	Glu	Ser	Val	
15					845					850					855	
	AGC	ACA	GAG	GAT	CTC	TCA	CCA	CCA	TCA	CCA	CCA	CTG	CCA	AAG	GAA	2934
					Leu											
20					860					865				-	870	
	AAT	TTA	AGA	GAA	GAG	GCA	TCA	GGA	GAC	CAA	AAA	TTA	CTC	AAC	ACA	2979
					Glu											
					875					880	-				885	
25	GGT	GAA	GGA	AAT	AAA	GAA	GCC	CCT	CTT	CAG	AAA	GTA	GGA	GCA	GAA	3024
	Gly	Glu	Gly	Asn	Lys	Glu	Ala	Pro	Leu	Gln	Lys	Val	Gly	Ala	Glu	
					890					895					900	
	GAG	GCA	GAT	GAG	AGC	CTA	CCT	GGT	CTT	GCT	GCT	TAA	ATC	AAC	GAA	3069
30	Glu	Ala	Asp	Glu	Ser	Leu	Pro	Gly	Leu	Ala	Ala	Asn	Ile	Asn	Glu	
					905					910					915	
	TCT	ACC	CAT	ATT	TCA	TCC	TCT	GGA	CAA	AAC	TTG	AAT	ACG	CCA	GAG	3114
	Ser	Thr	His	Ile	Ser	Ser	Ser	Gly	Gln	Asn	Leu	Asn	Thr	Pro	Glu	
35					920					925					930	
					AAT											3159
	Gly	Glu	Thr	Leu	Asn	Gly	Lys	His	Gln	Thr	Asp	Ser	Ile	Val	Cys	
40			•		935					940					945	

Gln Glu

	GAA	A ATC	a AA	A ATC	GAC	: ACI	GAT	CAG	AA	ACA	AGA	GAC	AA1	CTC	AC:	3.204
	Gli	ı Met	Lys	Met	Asp	Thr	Asp	Glr	Ası	Thr	Arg	g Gli	Asr	Let	Thi	:
					950					955					960	
5	GGI	` ATA	LAA 1	TCA	ACA	GTI	GAA	GAA	CCA	GTT	TCA	CCA	ATG	CTI	ccc	3249
	Gly	Ile	Asr	Ser	Thr	Val	Glu	Glu	Pro	Val	Ser	Pro	Met	Leu	Pro	,
					965					970					975	
															313	
	CCT	TCA	GCA	GTA	GAA	GAA	CGT	' GAA	GCA	GTG	TCC	* אא	א כייי	CCN	~~~	3294
10	Pro	Ser	Ala	Val	Glu	Glu	Ara	Glu	Ala	Val	Sa-	7 150	. MCI	GLA	. CIG	3294
					980		3			985	261	гÃг	ınr	Ala		
										763					990	
	GCA	TCA	ССТ	CCT	GC-T	ארא	አ ምር	CCN	221							
	Ala	Ser	Pro	Pro	212	Th-	Mot	33a	GCA	AAT	GAG	TCT	CAG	GAA	ATT	3339
15					995	1111	Met	Ala	Ala			Ser	Gln	Glu	Ile	
					333					100	0				100	5
	CAT	CAA	C N M	<i>~</i>												
	DAI.	GAA	GAI	GAA	GGC	ATC	CAC	AGC	CAT	GAA	GGA	AGT	GAC	CTA	AGT	3384
	Asp	GIU	Asp	GIU			His	Ser	His			Ser	Asp	Leu	Ser	
20					101	0				1015	5				102	0
20																
	GAC	AAC	ATG	TCA	GAG	GGT	AGT	GAT	GAT	TCT	GGA	TTG	CAT	GGG	GCT	3429
	Asp	Asn	Met	Ser	Glu	Gly	Ser	Asp	Asp	Ser	Gly	Leu	His	Gly	Ala	
					1025					1030					1035	5
25	CGG	CCA	GTT	CCA	CAA	GAA	TCT	AGC	AGA	AAA	AAT	GCA	AAG	GAA	GCC	3474
	Arg	Pro	Val	Pro	Gln	Glu	Ser	Ser	Arg	Lys	Asn	Ala	Lvs	Glu	Ala	• - / •
					1040					1045					1050	`
															1050	•
	TTG	GCA	GTC	AAA	GCG	GCT	AAG	GGA	GAT	بلملسل	بلبيلت	тст	እጥር	الميت	±rC±r	3519
30	Leu	Ala	Val	Lys	Ala	Ala	Lvs	Gly	Asn	Dhe	Val	Cvc	TIA	Dho	101	3213
				•	1055		-1-	,	щ	1060		Cys	116	PHE		
										1000					1065	•
	GAT	CGT	ጥርጥ	TTC	AC A	ממ	GGA	222	C 3 TD	m r 0						3564
	ASD	Ara	Ser	Dhe	Ara	Tara	Cl	AAA	GAI	TAC	AGC	AAA -	CAC	CTC	AAT	3564
35				1			GIY	Lys	Asp			Lys	Hıs	Leu	Asn	
-					1070	1				1075					1080	1
	ccc	~~~	mm-c					_								
	200	CAT	TTG	GTT	AAT	GTG	TAC	TAT	CTT	GAA	GAA	GCA	GCT	CAA	GGG	3609
	Arg	Hls	Leu	۷a:l			Туг	Tyr	Leu	Glu	Glu	Ala	Ala	Gln	Gly	
40					1085					1090					1095	
40																
	CAG	GAG	TAAT	G AA	ACTT	TGAA	CAA	.GGTT	TCA	GTTC'	TTAG	TT				3650

	TGTAAGGTAT	ATTACATTTT	ATATTCATTT	ATGATAGCAG	ACAACCTTTT	3700
	AAGATTGCTT	TAATTAGTAT	CTGATGTTGA	TTTTTAAGTG	GCATTCTTTT	3750
5	CCTTAGGACT	TTTTATGTAT	ACCTGTTGAT	TGTTGTGTAA	ATTTTAGTAA	3800
	ATCTAAGAGA	GTGTACTAAA	CCAGCAGGTA	TCTGTTAGCT	TATGTGTTTA	3850
10	ATTGAAATTA	GAAGGCTAAG	ATGGTATAAC	AGCATTTTAT	TGCTTTGTCC	3900
	AGCTACAACA	TGTCATTTTT	TTCTCCATGT	CTTATCTTCC	TGTTTCACTT	3950
	TAGTTTATTC	TTCGTTTTTT	ATTGAGATCT	ATAAAAAATT	GGCTTACTTA	4000
15	ATAGCAAATT	ACTTGAAGAA	TTTGCCTGCT	TTATATAAAG	TTAGCACTTT	4050
	AAGATTTTTT	TTTTAGAGAT	GAGAAGACAT	TTAAATTGAA	GAAAAATTCC	4100
20	CCCAGCAATA	GACAGTCTAT	CAGTCCAAGT	ATTTACTTCC	TGAGTTTTGA	4150
	TCAATATTTT	TTATTTGTGT	ATGTTAATCG	TCATAAAAAC	AGTGATTTTG	4200
	GTGTGTTTTT	TATTTTGGTG	CTTTAATGGC	TTAAGATGTT	GCACATTTTT	4250
25	TTTTTCTTTT	GGTTTCTGTT	TATGTTTTT	TGCCTATGCA	GTTAAATTTT	4300
	TCCTAGAAAT	AGCATTTGTG	TTGAACAGTA	ACACTTTATA	CATATATATA	4350
30	TGCATGTTTA	TTTTGTTTGG	CGTCTTTGGA	GGGATGCTTT	TAGACTTGTT	4400
	TGCAAAAGGG	CAGTTTTCTT	TTTCTTTGCT	GCAGTTGTCT	ATTTTGCAGA	4450
	ATAATAGTGT	GTGCAAGTTT	GTGAGCAAAT	GAAATATGCA	GGTTCAATCT	4500
35	ATTGATTTTG	ATTTTTACAT	СТТАТАТСТА	TGCCAGAATC	TGTATTTCAT	4550
	ATAACTTATT	TATTTCGAAT	GGATGTAGTA	AATTCACAGC	TATCAGTTTT	4600
40	GATTTTGCAA	TAAATAAACC	ACTAGGTTGC	ATGTCGAACA	AATTTTTATC	4650
	TCAAATACCA	ACCATCAGTT	TTTTTTTCA	TGTGTTTTGG	TACAGCTAAT	4700

	TCCTAATTG	I AGAGTGTTAJ	A ATGTTTGAG	G AGAACCTTT	I CTCATAGATG	4750
	GTTGGTGTT	ATATGGCNAC	TTTACAATA	A AGAGAACTG	T AAGTGATATT	4800
.5	TGGAAACTAC	AAACCTGGAA	TTAGGAGAT	A TAATTATTCO	TTCAAGTTTT	4850
	ATAGATATCA	CTTGGGAGAT	TCCAAAGCC	TAGCTATTAC	GCNGCAAACC	4900
10	TAGGATAAGA	AAGGTAGTAI	GAGTGCTGG1	AGACCAGCT	CAACATTTCC	4950
	TATATCAGAT	GAAAAAGGCT	GGTGAAACAA	GTACAGTCCA	GATTTTTAA	5000
	AATCATACTT	· TCTCAGGGAT	CTCCACAAAC	TGGTGGGTGT	CCTGGCTGTC	5050
15	TGTGTGATAG	CCTCTTTCTA	TAGGTGAGGC	CTCAAATGAA	TTGCAGCTAT	5100
	CCTGGTGTTC	CTATGAGGGC	ACTTGTATGA	AAAAGGCAGT	ACTCCAAAAC	5150
20	ATTTTTGATG	GTTCTTTGGC	CAGTTGCCAA	AGAGTGTGAA	AGAATCCAAT	5200
	AGAGGATTTT	TCTTACTGAT	AGCAGTCATT	CATTGCAGTA	AAATAAAATA	5250
	TGAATTCCCA	TTAGGGAATC	TTGAATTCTG	ACCTCCCATA	CTCCGTTTTG	5300
25	AAATAACCAC	TTATATTTCA	TTTTTTAAAA	ATCTGATGAT	CTCTTTGAGG	5350
	CAGGTTTCAG	ATTTGGCAGT	ACAACATGAA	AGATTAGGAA	AAGCATTAAT	5400
30	AACGTGTGGG	TGGAAAGCTT	GTTAAAAATC	TGAGAGTGAA	GTTTGAGTTA	5450
50	AAAGTTGTTT	GACATGGCAT	TGACTGGGAG	GCCAAAGATT	TAAAGAAGCG	5500
	GAAGATTCTT	CTCTTAAGAC	ATGAGGAGTA	AGTTGTGTGA	TAATGGTATG	5550
35	TGTTTTGTGT	GCATGAATGĢ	ACATTGTAAA	TGTTGAATTC	TAGGCTCCGA	5600
	CAATCATTGT	CAACAGAAGA	TAAAGCTGCA	AATATTTATG	TTTTAAAA	5648

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 756 base pairs
 - (B) TYPE: nucleic acid
- :5 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 10 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
- 15 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
- 20 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
- 25 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:2:FROM 1 TO 756
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGT AAG CCA TGC CAA

15

30 Cys Lys Pro Cys Gln

165

TAT GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT 60

Tyr Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val

170 175 180

CAC AGT GCT AAG AAA TTT TTT GTG GAA GAG AGT GCA GAG AAG CAG 105 His Ser Ala; Lys Lys Phe Phe Val Glu Glu Ser Ala Glu Lys Gln 185 190 195

	GC	'A A	AA .G	CC /	\GG	GA/	A TC	T G	C T	CT T	CC A	CT	GC	A GA	A G	AG.	CCA	GAT	150
	Al	a L	ys A	la /	lrg	Gli	ı Se	r Gl	y Se	er s	er T	'hr	Ala	Gl	ц G)	l 11	G) v	Asp	
						.200)				2	05						210	
	TT	C T	CC A	AG (GC	CCC	AT	r cg	CT	T G	AC C	GC	TGC	GG	C TZ	۸C /	AAT	יים עי	105
4	Ph	e Se	er L	y:s G	ly	Pro) Ile	e Ar	g Cy	s As	p A	rg	Cys	G1	v Tv	n d	Asn	Thr	
						:215	;					20	•		•			225	
	AA'	T CO	SA T	AT G	AT	CAC	TAT	r Ac	A GC	A CA	rc c	TG	AAA	CA	C CA	.c z	ACC	AGA	240
10	AS	n Aı	g T	YY A	sp	His	Туз	Th	r Al	a Hi	s L	eu	Lys	His	s Hi	s :	Phr	Arg	
10	l					230						35						240	
	0.01																-		
	B.3.	ا الحاد	G GZ	AT A	AT	GAG	CGA	GT	TA	C AA	G T	GT	ATC	ATT	TG	C A	CA	TAC	285
	Ali	1 61	у Аз	SP A	sn	Glu	Arg	Va.	l Ty	r Ly	s Cy	/5	Ile	Ile	Cy.	s I	hr	Tyr	
15						245					25							255	
	ACI	י אר	א כיז	אל בי	~~	~~~													
	Thr	. Th	r Va	l e	3C 1	Clu	TAT	CAC	TG:	G AG	G AA	LA I	CAT	TTA	AG	A A	AC	CAT	330
				.1 5	=1 '	260	Tyr	His	Tr	o Ar			His	Leu	Arg	g A	sn	His	
					•	200					26	5						270	
20	TTT	. cc	A AG	G AJ	AA (GTA	TAC	ACZ	TC	r GG									
	Phe	Pr	o Ar	g Ly	s 1	Val	Tvr	Thr	· Cvs	Gl ₂	, T.	A :	TGC	AAC	TAT	r T	TT	TCA	375
				•	2	275	-1-		-y.	, GI	7 Ly 28		_ys	Asn	Туг	. P.			
											20	•						285	
	GAC	AG	AA A	A AA	C A	TAL	TAT	GTT	CAG	CAT	GT	T A	4GA	ידי) מ	רמיז	. א	~	CCX	420
25	Asp	Arg	J Ly	s As	n A	sn	Tyr	Val	Glr	His	: Va	 1 &	\ra	Thr	Hic	יים.	om (31	420
					2	90					29		5					300	
	GAA	CG	c cc	AT A	TA	AA	TGT	GAA	CTT	TGI	. cc.	гт	AC	TCA	AGT	TO	T (CAG	465
20	Glu	Arg	y Pro	ту	r L	ys	Cys	Glu	Leu	Cys	Pro	o T	уr	Ser	Ser	S€	er (Sln	
30					3	05					310							315	
	AAG	ACT	CAT	CT.	A A	CT .	AGA	CAT	ATG	CGT	ACT	r c	AT '	TCA	GGT	GA	G A	LA G	510
	Lys	Thr	His	Le			Arg	His	Met	Arg	Thi	: н	is :	Ser	Gly	Gl	u I	ys	
35					3	20					325	5					3	30	
	CCD	יוייטיט		. TC:	. ~		^												
	CCA	Dhe	TAAA	16	r G.	AT (CAG	TGC	AGT	TAT	GTG	G	CC 7	rct	TAA	CA	A C	TA:	555
	Pro	rne	Буб				σIN	Cys	ser	Tyr			la s	Ser	Asn	Gl	n H	is	
					. د	35					340)					3	45	
40	GAA	GTA	ACC	רפי	י כי	ልጥ ሰ	ברא	א כי א	C	~~~	0 5 ~								
	GAA Glu	Va)	Thr	Arc	ים ב	ie 7	aca.	nun n	CAG	GIT	CAC	: A <i>l</i>	AT G	GG	CCT	AA.	A C	CT	600
	Glu				, n. 3!		ua.	ar 9	OIII	val			sn G	īУ	Pro	Ly.			
					٠.						355						3	60°	

- 42 -

CTT AAT TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC 645 Leu Asn Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn 365 370 375

5 TTC AAA AAA CAT GTA GAG CTA CAT GTG AAC CCA CGG CAG TTC AAT 690
Phe Lys Lys His Val Glu Leu His Val Asn Pro Arg Gln Phe Asn
380 385 390

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG 735

10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln

395 400 405

TAT CAC TTC AAA TCT AAG CAT
Tyr His Phe Lys Ser Lys His
15 410

- (2) INFORMATION FOR SEQ ID NO: 3:
- (i) SEQUENCE CHARACTERISTICS
- 20 (A) LENGTH: 1407 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
- 25 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
- 30 (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
- 35 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
- 40 (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995

(X) RELEVANT RESIDUES IN SEQ ID NO:3:FROM 1 TO 1407 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	G J	ATG	GCA (GAA												10
:5	t	Met .	Ala (Glu												10
				75												
	CTO	ATC	. ככנ	בייתום ב	r ccc	י כאיז										
	I.e.i	Met	- Dr	. 1/2]		GAL	AAL	AAC	_ TTF	r rcz	A GAT	r AGI	GAJ	A GA	A GGA	55
10				, val	. 61)	ASP) ASI	ASI	1 Phe	e Ser	r Ası	Ser	Gli	Gl	ı Gly	
.0					80					85					90	
	~													•		
	GAA	GGA	CTI	GAA	GAG	TCT	GCI	GAT	ATA	AAA	A GGI	GAA	CCI	CAT	GGA	100
	Glu	Gly	/ Let	Glu	Glu	Ser	Ala	Asp	Ile	Lys	Gly	Glu	Pro	His	Gly	
					95					100					105	
15						•										
	CTG	GAA	AAC	ATG	GAA	CTG	AGA	AGT	TTG	GAA	CTC	AGC	GTO	GTE	GAA	145
	Leu	Glu	Asn	Met	Glu	Leu	Arg	Ser	Leu	Glu	Leu	Ser	Val	Val	Glu	143
					110		_			115			• • • • • • • • • • • • • • • • • • • •	VEI	120	
											,				120	
20	CCT	CAG	CCT	GTA	TTT	GAG	GCA	TCA	GGT	GCT	CCA	Cam	N COOL	m> 0		
	Pro	Gln	Pro	Val	Phe	Glu	Ala	Ser	Gly	בוש	Pro	GAI	All	TAC	AGT	190
					125			562	Gry			Asp	ile	Tyr		
										130					135	
	TCA	AAT	444	CCT	حبي	GCC	COTT	63.3								
25	Ser	Acn	Tare	אות	T	33-	-	GAA	ACA	CCT	GGA	GCG	GAG	GAC	AAA	235
	001	N3II	Dys	Ald		AIA	Pro	Glu	Thr	Pro	Gly	Ala	Glu	Asp	Lys	
					140					145			•		150	
	000													•		
	GGC	AAG	AGC	TCG	AAG	ACC	AAA	CCC	TTT	CGC	TGT	AAG	CCA	TGC	CAA	280
20	Gly	Lys	Ser	Ser	Lys	Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	
30					155					160					165	
	TAT	GAA	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC	ATC	AGA	GTT	325
	Tyr	Glu	Ala	Glu	Ser	Glu	Glu	Gln	Phe	Val	His	His	Ile	Ara	Val	
					170					175				9	180	
35															100	
	CAC	AGT	GCT	AAG	AAA	TTT	للللل	GTG	ממם	GAC	AGT	CCN	C3.0		~~~	
	His	Ser	Ala	Lvs	Lvs	Dhe	Dhe	1/21	Clu	Clu	Ser	GCA	GAG	AAG	CAG	370
				-,-	185		- 110	Val	GIU		Ser	AIA	GIu	Lys	Gln	
				i i	100					190					195	
40	GC ₂	תתת	CCC	700	C 2 2	ma-										
40	DI-	7	N1 -	AGG	GAA	TCT	GGC	TCT	TCC	ACT	GCA	GAA	GAG	GGA	GAT	415
	wrg	ьys	АТА	Arg		Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
					200					205					210	

	TTC	TCC	AAG	GGC	CCC	. אדר ב	CGC	ידטע.	GNC	CGC	-שרביר	CCC	.ms ~	· R R T	ACT	450
						Ile										-460
			-,-	-0-1	215	.110	AL.9	Cys	ASp		Cys	GIY	ıyr	Asn		
					227					220					225	
·5	AAT	CGA	TAT	GAT	CAC	TAT	ACA	GCA	CAC	CTG	AAA	CAC	CAC	ACC	AGA	-505
						Тут										203
					230	•				.235	-,-				240	
															240	
	GCT	GGG	GAT	AAT	GAG	CGA	GTC	TAC	AAG	TGT	ATC	ATT	TGC	ACA	TAC	550
10						Arg										
					245				-	250			4)		255	
	ACA	ACA	GTG	AGC	GAG	TAT	CAC	TGG	AGG	AAA	CAT	TTA	AGA	AAC	CAT	595
	Thr	Thr	Val	Ser	Glu	Tyr	His	Trp	Arg	Lys	His	Leu	Arg	Asn	His	
15					260					265					270	
	TTT	CCA	AGG	AAA	GTA	TAC	ACA	TGT	GGA	AAA	TGC	AAC	TAT	TTT	TCA	640
	Phe	Pro	Arg	Lys	Val	Tyr	Thr	Cys	Gly	Lys	Cys	Asn	Tyr	Phe	Ser	
					275					280					285	
20																
						TAT										685
	Asp	Arg	Lys	Asn	Asn	Tyr	Val	Gln	His	Val	Arg	Thr	His	Thr	Gly	
					290					295					300	
3.5																
25						TGT										730
	Glu	Arg	Pro	Tyr		Cys	Glu	Leu	Cys	Pro	Tyr	Ser	Ser	Ser	Gln	
					305					310					315	
						_										
20						AGA										775
30	гÀг	Thr	HIS	Leu		Arg	His	Met	Arg		His	Ser	Gly	Glu	Lys	
					320					325					330	
	CCN	ملحلحك		~~~	C	C) C	maa	1 0 m								
						CAG										820
35	FIU	FILE	Dys	Cys		Gln	Cys	ser	Tyr		Ата	ser	ASI	GIN		
					335					340					345	
	C A A	СТЪ	אככ	CGC	ር አ ጥ	GCA	ארא	C	COTO	C C	7. F. M	~~~				0.55
																865
	424	- 41		y	350	Ala	Arg	, G111	val		ASII	стА	Pro	гÀг		
10					J J U					355					360	

	CT	ı AA	T TG	כ ככ	A CA	C TG	GA:	AT T	C AA	A AC	A GC	A GA	r .AG	ARG	CAA	910
	Lei	ı Ası	п Су	s Pro) Hi	s Cys	Asp	э Туз	Ly:	s Th	r Ala	a Asp	Arc	.Se	r Ası	n
					36	5				370		-	-	,	379	
5	TTC	LAA	A AA	A CAT	GT)	A GAG	CTA	CAT	GTO	AA	C C C 2	A CGC	CAG	السال :	רממ ר	r 955
	Phe	Lys	Ly	s His	Va]	Glu	Leu	His	Va]	l Asr	n Pro	Arg	Glm	Phe) Acr	, ,,,,
					380)				385		_			390	
															330	,
	TGC	CCI	GT	A TGI	GAC	TAT	GCA	GCT	TCC	: AAG	AAG	TGT	AAT	נידים י	י ראכ	1000
10	Cys	Pro	Va]	Cys	Asp	Tyr	Ala	Ala	Ser	Lys	Lvs	Cys	Asn	T.et	Gla	1000
										•	•	-,-		-	. 011	٠.
					395					400					405	
	TAT	CAC	TTC	AAA:	TCI	AAG	CAT	CCT	ACT	TGI	. cci	` AAT	444	מיא	בטב מתמ	1045
	Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Cys	Pro	Asn	Lvs	Thr	Met	1045
15					410					415			-,-		420	
	GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	GCT	1090
	Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	1030
					425					430		•			435	
20																
	GAC	TTG	CCT	GAT	AAT	ATT	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	1135
	Asp	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
					440					445					450	
25	ACA	AAA	ATA	AAA	GGG	GAT	GTG	GCT	GGA	AAG	AAA	AAT	GAA	AAG	TCC	1180
	Thr	Lys	Ile	Lys	Gly	Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
					455					460				-	465	
20	GTC	AAA	GCA	GAG	AAA	AGA	GAT	GTC	TCA	AAA	GAG	AAA	AAG	CCT	TCT	1225
30	Val	Lys	Ala	Glu	Lys	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
					470					475					480	
						•										
	AAT	TAA	GTG	TCA	GTG	ATC	CAG	GTG	ACT	ACC	AGA	ACT	CGA	AAA	TCA	1270
3.5	Asn	Asn	Val	Ser	Val	Ile	Gln	Val	Thr	Thr	Arg	Thr	Arg	Lys	Ser	
35					485					490					495	
	GTA	ACA	GAG	GTG	AAA	GAG	ATG	GAT	GTG	CAT	ACA	GGA .	AGC .	TAA	TCA	1315
	Val	Thr	Glu	Val	Lys	Glu	Met	Asp	Val	His	Thr	Gly	Ser 2	Asn	Ser	
4.0					500					505					510	
40																

- 46 -

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1360 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 1405 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser 530 535 540

AC

1407

- (2) INFORMATION FOR SEO ID NO: 4:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 1090 base pairs
 - (B) TYPE: nucleic acid
- 15 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 20 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
- 25 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
- 30 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
- 35 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:4:FROM 1 TO 1090
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

	С	AAG Lys	GGC Gly	Pro	ATT Tle	CGC Arg	TGT Cys	GAC Asp	.CGC Arg	TGC Cys	GGC Gly	TAC	AAT Asn	ACT		-40
				215					220					225		
5	מ מ	T CC	·> m>	.		_										
-	λc	ח אר	M 1A	II GA	T CA	C TA	T AC	A GC	A CA	C CI	G AZ	VA CA	C C	C AC	C AG	85
			9 1	ı As	р ні 23	.в ту •	T Th	r Al	a Hi	s Le	u Ly	's Hi	s Hi	s Th	r Arg	J
					23	U				23	5				240)
	GC	T GG	G GA	T AA	T GA	e ce	ש ביי	C TIN	~						A TAC	
10	Ala	a Gl	y As	p As:	n Gl	u Ar	a Va	יים או אור ביים	- AA	G TG	T AT	C AT	T TG	C AC	A TAC r Tyr	130
				_	24	5	J • W.	y.	L Lly	s cy 25		e II	е су	s Th		
										23	O				255	
	AC	A. AC	A GT	G AG	GA	G TA	T CAC	TGO	AG	G AA	A CA	ئىلىش بال	מ מ	מת מ	C CAT	125
	Thi	Th	r Va	l Ser	Gl:	и Туг	r His	Tr	Arc	T Lv	s Hi	s Lei	ı Arı	- A	n His	175
15					26	0		_		26				y As	270	
	TTI	CC	A AGO	AA.	GT	A TAC	ACA	TGI	GG	LAA A	A TG	C AAC	TA	r Trr	TCA	220
	Phe	Pro	Arg	J Lys	Va]	Tyr	Thr	Cys	Gly	/ Lys	су:	s Ası	тул	Phe	Ser	
20					275	5				280					285	
20	G N C	י ארי														
	Acn	AGA	. AAA	AAC	TAA	TAT	GTT	' CAG	CAI	GTI	AGA	ACI	CAI	ACA	GGA	265
	رود	. AL 5	בענו	ASI	AST	Tyr	Val	Gln	His			J Thr	His	Thr	Gly	
					290	,				295					300	
25	GAA	CGC	CCA	TAT	מממ	TCT	ר מא א	COM	man							
	Glu	Arg	Pro	Tvr	Lvs	Cve	GAA Glu	Lan	Corn	CCT	TAC	TCA	AGT	TCT	CAG	310
					305	-,2	Olu	Deu	Cys	310	тут	Ser	Ser	Ser		
										310					315	
	AAG	ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	АСТ	רבת	י דרא	GCT	CAC	770	355
30	Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thr	His	Ser	GUV	GAG	Tare	355
					320				J	325			,	014	330	
															330	
	CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	400
	Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
35					335					340					345	
	GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	445
	GIU	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lys	Pro	
40					350					355					360	
70																

	CTT	TAA	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	490
			Cys													
					365					370					375	
·5	TTC	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	535
	Phe	Lys	Lys	His	Val	Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
					380					385					390	
			GTA													580
10	Cys	Pro	Val	Cys	Asp	Tyr	Ala	Ala	Ser	Lys	Lys	Cys	Asn	Leu	Gln	
													•			
					395					400					405	
			TTC													625
	Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Cys	Pro	Asn	Lys	Thr	Met	
15					410					415					420	
																670
	Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	
••					425					430					435	
20																
			CCT													715
	Asp	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
					440					445					450	
25																
25			ATA													760
	Thr	Lys	Ile	Lys		Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
					455					460					465	
30															TCT	805
30	vaı	Lys	Ala	GIU		Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
					470					475					480	
			GTG													850
25	Asn	Asn	Val	Ser		He	Gin	Val	Thr		Arg	Thr	Arg	Lys	Ser	
35					485					490					495	
	O.T. >															
			GAG													895
	val	Inr	Glų	val		Glu	Met	Asp	Val		Thr	Gly	Ser	Asn	Ser	
4 ∩					500					505					510	
411																

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 940 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 985
Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser
530 535 540

ACA AAA AAG AAA AAG AAG GTA GAA AGC AAA TCC AAA AAT AAT AGT 1030

Thr Lys Lys Lys Lys Val Glu Ser Lys Ser Lys Asn Asn Ser

545

550

555

CAG GAA GTG CCA AAG GGT GAC AGC AAA GTG GAG GAG AAT AAA AAG 1075 Gln Glu Val Pro Lys Gly Asp Ser Lys Val Glu Glu Asn Lys Lys 560 565 570

CAA AAT ACT TGC ATG Gln Asn Thr Cys Met

1090

20

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 928 base pairs
 - (B) TYPE: nucleic acid
- 25 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 30 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
- 35 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts 40 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80

		\ _ / .	1330	. مته												
		(F)	PAGE	S:												
		(G)	DATE	: Ma	rch :	24,	1995									
		(K)	RELE	TMAV	RES	IDUE	S IN	SEO	ID	NO : 5	: FRO	м 1	TO 9	28		
·5				CE D								_				
	CA (GCA -	CAC	CTG :	AAA (CAC	CAC .	ACC .	AGA							26
				Leu :												20
		235			-		240		5							
10																
	GCT	GGG	GAT	AAT	GAG	CGA	GTC	TAC	AAG	TGT	ATC	ידינא	TCC	אכא	TAC	71
				Asn												, ,
			_		245	_		- 2 -	-,-	250			Cyb		255	
															233	
15	ACA	ACA	GTG	AGC	GAG	TAT	CAC	TGG	AGG	AAA	CAT	TTA	AGA	AAC	СУТ	116
				Ser												
					260	•				265					270	
															2.0	
	TTT	CCA	AGG	AAA	GTA	TAC	ACA	TGT	GGA	AAA	TGC	AAC	TAT	Juluk	TCA	161
20				Lys												
					275	-		•	•	280	-3-		-1-		285	
	GAC	AGA	AAA	AAC	AAT	TAT	GTT	CAG	CAT	GTT	AGA	ACT	CAT	ACA	GGA	206
				Asn												
25					290					295	•				300	
	GAA	CGC	CCA	TAT	AAA	TGT	GAA	CTT	TGT	CCT	TAC	TCA	AGT	TCT	CAG	251
				Tyr												
					305					310	_				315	
30																
	AAG	ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	ACT	CAT	TCA	GGT	GAG	AAG	296
	Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thr	His	Ser	Gly	Glu	Lys	
					320					325			_		330	
35	CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	341
	Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
					335					340					345	
			·													
	GAA	GTA		CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	386
40				Arg												
					350		_			355		-		-	360	

	<u></u>	מו ידי	TO TO			_										
	· · · · · · · · · · · · · · · · · · ·	1 AA	.1 10	3C C(CA C	AC TG	T GA	TA	CAA	A AC	A GC	A GA	T AG	A AG	C AAC	431
	.he	u :As	n Cy	/s Pi	ro Hi	rs Ca	s As	p Ty	r Ly	s Th	r Ala	a As	o Ar	g Se:	r Asn	
					36	55				37					375	
<i>'</i> 5	TT	- AA	AA A	A CA	AT GI	'A GA	G CT	A CAT	r GT	AA E	ב רכז	٠	ב ראי	C HAND	C AAT	450
	Phe	Ly	s .Ly	s Hi	s Va	l Gl	u Lei	ı His	va.	l Agr) Pro	. CG(r Cl	2 II(AA1 Asn	.476
					38	0				385		, 11.5	, GI	.i Pile		
										-	•				390	
	TGC	cc:	T GT	A TG	T GA	C TA	T GC	GCI	TCC	. AAG	AAC	י דיניז	ית מי	מידורים ז	CAG	531
10	Cys	Pro	va	1 су	s As	p Ty:	r Ala	Ala	Ser	Lvs	Lve	Cve	AA.	. CIA	CAG	521
										-,-	. - , 5	Cys	, war	Tien	GIN	
					39	_				400)				405	
	TAT	CAC	TT	C AA	A TC	T AAC	CAT	CCT	ACT	TGT	CCT	ם אבר	מממי	מים	NTC.	566
	Tyr	His	Ph	e Ly	s Se	r Lys	His	Pro	Thr	Cvs	Pro	Asn	Lve	Thr	Mor	566
15					41	0				415			2 ,5	1111	420	
	GAT	GTO	TC	AA A	A GT	AAA E	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	CCT	611
	Asp	Val	Se	r Ly	s Va	l Lys	Leu	Lys	Lys	Thr	Lvs	Lvs	Ara	Glu	Δla	911
					42	5			_	430	-3-	-1-	9	O1u	435	
20																
	GAC	TTG	CC	GA:	CAA 1	TTA 1	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	656
	Asp	Leu	Pro	As _I	Asr	lle	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	0.50
					440)				445					450	
25	ACA	AAA	ATA	AAA A	GGG	GAT	GTG	GCT	GGA	AAG	AAA	AAT	GAA	AAG	TCC	701
	Thr	Lys	Ile	Lys	Gly	' Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
					455					460				•	465	
20	GTC	AAA	GCA	GAG	AAA	AGA	GAT	GTC	TCA	AAA	GAG	AAA	AAG	CCT	TCT	746
30	Val	Lys	Ala	Glu	Lys	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
					470					475			-		480	
							•									
	AAT	TAA	GTG	TCA	GTG	ATC	CAG	GTG	ACT	ACC	AGA .	ACT	CGA	AAA	TCA	791
	Asn	Asn	Val	Ser	Val	Ile	Gln	Val	Thr	Thr	Arg	Thr .	Arg	Lys .	Ser	
35					485					490	_		•		495	
	GTA	ACA	GAG	GTG	AAA	GAG	ATG	GAT	GTG	CAT .	ACA (GGA 2	AGC	AAT '	TCA	836
	Val	Thr	Glu	Val	Lys	Glu	Met	Asp	Val :	His '	Thr (Gly :	Ser	Asn !	Ser	
				•	500					505					510	
40														•		

- :52 --

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 881 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val 515 520 525

5 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TCT TCA 926
Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu Ser Ser
530 540

AC

928

10

- (2) INFORMATION FOR SEQ ID NO: 6:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 1791 base pairs
- 15 (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
- 20 (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
- 25 (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- 30 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
- 35 (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:6:FROM 1 TO 1791
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
- 40 CACCCTCTGC AGCCCCACTC CTGGGCCTTC TTGGTCCACG ACGGCCCCAG

50

100

CACCCAACTT TACCACCCTC CCCCACCTCT CCCCCGAAAC TCCAGCAACA

	AAC	LAAAE	AGTA	GTC	GAGA	AG G	SAGCG	GCG/	AC T	CAGG	GTCG	2 000	GCCC(CTCC		.150
	TCA	ACCG#	\GGA	AGG	CGAA	TA C	AGTI	•								175
.5	ATG Met	GCC Ala	: ACC	CAG	GTA Val	ATG Met	GGG Gly	CAG	TC:	TCT Ser	r GG/	A GGA	A GG/	A GGC	G CTG / Leu 15	.220
10	TTT	he T	AGC	AGT Ser S	GGC er G	AAC ly A	ATT sn I	GGA le G	ATG	GCC let A	CTG	CCT	AAC	GAC	ATG	265 et
15	TAT	GAC Asp	TTG Leu	CAT His	GAC Asp 35	CTT Leu	TCC	AAA Lys	GCI Ala	' GAA	CTG	GCC Ala	GCA	CCI Pro	30 CAG Gln 45	310
20	CTT Leu	ATT Ile	ATG Met	CTG	GCA Ala 50	AAT Asn	GTG Val	GCC Ala	TTA Leu	ACT Thr	GGG Gly	GAA Glu	GTA Val	AAT Asn	GGC Gly	355
,	AGC Ser	TGC Cys	TGT Cys	GAT Asp	TAC Tyr 65	CTG Leu	GTC Val	GGT Gly	GAA Glu	GAA Glu 70	AGA Arg	CAG Gln	ATG Met	GCA Ala	GAA Glu 75	400
25	CTG Leu	ATG Met	CCG Pro	GTT Val	GGG Gly 80	GAT Asp	AAC Asn	AAC Asn	TTT Phe	TCA Ser 85	GAT Asp	AGT Ser	GAA Glu	GAA Glu	GGA Gly 90	445
30	GAA Glu	GGA Gly	CTT Leu	GAA Glu	GAG Glu 95	TCT Ser	GCT Ala	GAT Asp	ATA Ile	AAA Lys 100	GGT Gly	GAA Glu	CCT Pro	CAT His	GGA Gly 105	490
35	CTG Leu	GAA Glu	AAC Asn	ATG Met	GAA Glu 110	CTG Leu	AGA Arg	AGT Ser	TTG Leu	GAA Glu 115	CTC Leu	AGC Ser	GTC Val	GTA Val	GAA Glu 120	535
40	CCT Pro	CAG Gln	CCT Pro	GTA Val	TTT Phe 125	GAG Glu	GCA Ala	TCA Ser	GGT Gly	GCT Ala 130	CCA Pro	GAT Asp	ATT Ile	TAC Tyr	AGT Ser 135	580

	TCA	TAA	AAA	GCT	CII	GCC	CCI	GAA	ACA	CCT	GGA	GCG	GAG	GAC	AAA	625
											Gly					
					140					.145					150	
'5	GGC	AAG	AGC	TCG	AAG	ACC	AAA	CCC	TTT	CGC	TGT	AAG	CCA	TGC	CAA	. 670
	Gly	Lys	Ser	Ser	Lys	Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	l
					155					160					165	
	TAT	GAA	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC	ATC	AGA	GTT	715
10	Tyr	Glu	Ala	Glu	Ser	Glu	Glu	Gln	Phe	Val	His	His	Ile	Arg	Val	
					170	•				175			•		180	
											AGT					
	His	Ser	Ala	Lys	Lys	Phe	Phe	Val	Glu	Glu	Ser	Ala	Glu	Lys	Gln	
15					185					190					195	
											GCA					
	Ala	Lys	Ala	Arg	Glu	Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
					200					205					210	
20																
											TGC					850
	Phe	Ser	Lys	Gly	Pro	Ile	Arg	Cys	Asp	Arg	Cys	Gly	Tyr	Asn	Thr	
					215					220					225	
25											AAA					895
	Asn	Arg	Tyr	Asp	His	Tyr	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	
					230					235					240	
••											ATC					940
30	Ala	Gly	Asp	Asn	Glu	Arg	Val	Tyr	Lys	Cys	Ile	Ile	Cys	Thr	Tyr	
					245					250					255	
											CAT					985
	Thr	Thr	Val	Ser	Glu	Tyr	His	Trp	Arg	Lys	His	Leu	Arg	Asn	His	
35					260					265					270	
																1030
	Phe	Pro	Arg;	Lys	Val	Tyr	Thr	Cys	Gly	Lys	Cys	Asn	Tyr	Phe	Ser	
					275					280					285	
40																

	GA	AG2	AA A	A AA	C AA!	TAT	GT	CA(G CA	r gr	T AG	A :AC	r ca	ר :בר	A GG	A 1075
	Asj	Arg	J Ly	s As	n Ası	Tyz	Va]	Gli	n His	s Va	l Arg	Th	Hi	- "ምር ዓ. "ምክ	r G1:	M 10/5
					290)				.29				3 111		
															30	J
5	GAZ	CGC	cc	A TA	LAA T	TGI	GA.	CTT	TGT	r cc.	T TAC	יידרי י	א א	r -m~	ר כאנ	G 1120
	Glu	Arg	J Pro	о Туз	Lys	Cys	Glu	Let	ı Cys	Pro	э Туг	Ser	Set	- 50	r (2) -	1120
					305	;			•	31				. 56.	315	
											-				313	•
	AAG	ACI	CA	r cta	ACI	' AGA	CAT	ATG	CGI	' AC	CAT	TCA	GGT	י האר	. אמ	1165
10	Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thi	r His	Ser	เดา	COAC) Jac	
					320	ı				325			 -	GIL	330	
															336	
	CCA	TIT	' AAZ	TGI	GAT	CAG	TGC	AGT	TAT	GTO	GCC	TCT	דמב '	ממים י	ר מים	1210
	Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gla	Hic	1210
15					335				_	340					345	
	GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	ССТ	1255
	Glu	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lvs	Pro	1233
					350					355		•		-,-	360	
20																
	CTT	AAT	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	1300
	Leu	Asn	Cys	Pro	His	Cys	Asp	Tyr	Lys	Thr	Ala	Asp	Arg	Ser	Asn	
					365					370					375	
25																
25	TTC	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1345
	Pne	Lys	Lys	His	Val	Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
					380					385					390	
	-		_													
30	TGC	CCT	GTA	TGT	GAC	TAT	GCA	GCT	TCC	AAG	AAG	TGT	TAA	CTA	CAG	1390
30	Cys	Pro	Val	Cys	Asp	Tyr	Ala	Ala	Ser	Lys	Lys	Cys	Asn	Leu	Gln	
					395					400					405	
																
	TAT	CAC	TTC	AAA	TCT	AAG	CAT	CCT	ACT	TGT	CCT	AAT	AAA	ACA	ATG	1435
35	Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Суѕ	Pro	Asn	Lys	Thr	Met	
دد					410					415					420	
	63 -															
	GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	GCT	1480
	Asp	val	Ser			Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	
40				•	425					430					435	
4 U																

GAC	TIG	CCT	GAT	AAT	ATT	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	1525
qaA	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
				440					445					450	

- 5 ACA AAA ATA AAA GGG GAT GTG GCT GGA AAG AAA AAT GAA AAG TCC 1570 Thr Lys Ile Lys Gly Asp Val Ala Gly Lys Lys Asn Glu Lys Ser 455 460 465
- GTC AAA GCA GAG AAA AGA GAT GTC TCA AAA GAG AAA AAG CCT TCT 1615 10 Val Lys Ala Glu Lys Arg Asp Val Ser Lys Glu Lys Lys Pro Ser 470 480
- AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1660 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser 485 490 495
 - GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1705

 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser

 500 505 510

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1750
Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val
515 520 525

- 25 GAC AGC CAT TCT TTA CAT GGT CCT GTG AAT GAT GAG GAA TC 1791
 Asp Ser His Ser Leu His Gly Pro Val Asn Asp Glu Glu
 530 535
 - (2) INFORMATION FOR SEQ ID NO: 7:
- 30 (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 3705 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
- 40 (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA

										-						
	(x) PU	BLIC	ATIO	N IN	FORM	ATIO	N:								
									· Z	Tan	in-P		~		Tole	_
	Ara	al, .	Juan	, Zh	eng,	Ying	cond	a, B	outr	os.	mich.	amır ael	ez J	ose,	Tole chule	:do-
	Ye]	lena	·М.,	Fro	nman.	Mic	chael	l A.	Kr	aner	211	635	D	M3	el, G	:I,
5		(B)	TITI	Æ: 1	REST:	. A 4	4a mma	alia,		lene	, .ou.	san .	المارة المارة المارة المارة	mana	eı, G Restr	all
	Soc	lium	Char	mel	Gene	Ехт	Teg	zion	+0.1	Jon-	 	roce	ın t	nat -	Restr	icts
		(C)	JOUR	ENAI.	: Cel	. <u></u>	71686	31011	10 1	мент	ons					
			VOLU			•										
			ISSU													
10			PAGE													
		(G)	DATE	: Ma	rch	24	1995							•		
					RES				מד ו	NO.	7 . TTD C					
	(xi) SE	QUEN	CE I	ESCR	דידם.	ON .	SEO	TD N	10 · 7	FRU)M I	10 3	3705		
							.011.	סבפ	10 1	10:7	į					
15	GA	ACT	CGA	AAA	TCA											
			Arg													1
			_	-	495											
	GTA	ACA	GAG	GTG	AAA	GAG	ATG	GAT	GTG	CAT	י ארא	GGA	A.C.C	י אאי	TCA	-
20	Val	Thr	Glu	Val	Lys	Glu	Met	Asp	Val	His	Thr	G G J	Ser	. AAI	Ser	59
					500					505		Cly	361	ASII	510	
							•								310	
	GAA	AAA	TTC	AGT	AAA	ACT	AAG	AAA	AGC	AAA	AGG	AAG	CTG	CVV	GTT	104
	Glu	Lys	Phe	Ser	Lys	Thr	Lys	Lys	Ser	Lys	Arg	Lvs	Leu	Glu	Val	103
25					515			_		520	_	- 2 -		-	525	
															323	
	GAC	AGC	CAT	TCT	TTA	CAT	GGT	CCT	GTG	AAT	GAT	GAG	GAA	TCT	TCA	149
	Asp	Ser	His	Ser	Leu	His	Gly	Pro	Val	Asn	Asp	Glu	Glu	Ser	Ser	
					530					535					540	
30																
	ACA	AAA	AAG	AAA	AAG	AAG	GTA	GAA	AGC	AAA	TCC	AAA	AAT	AAT	AGT	194
					Lys											
					545					550		-			555	
															· -	
35	CAG	GAA	GTG	CCA	AAG	GGT	GAC	AGC	AAA	GTG	GAG	GAG	AAT	AAA	AAG	239
					Lys											
					560					565				•	570	

CAA AAT ACT TGC ATG AAA AAA AGT ACA AAG AAG AAA ACT CTG AAA 284

580

585

40 Gln Asn Thr Cys Met Lys Lys Ser Thr Lys Lys Lys Thr Leu Lys

	AAT	AAA	TCA	AGT	AAG	AAA	AGC	AGT	AAG	CCT	CCT	CAG	AAG	GAA	CCT	329
	Asn	Lys	Ser	Ser	Lys	Lys	Ser	Ser	Lys	Pro	Pro	Gln	Lys	Glu	Pro	
					590					59 5					600	
5			AAG													374
	Val	Glu	Lys	Gly	Ser	Ala	Gln	Met	Asp	Pro	Pro	Gln	Met	Gly	Pro	
					605					610					615	
			ACA													419
10	Ala	Pro	Thr	Glu	Ala	Val	Gln	Lys	Gly	Pro	Val	Gln	Val	Glu	Leu	
					620					625			-		630	
			ccc													464
	Pro	Pro	Pro	Met	Glu	His	Ala	Gln	Met	Glu	Gly	Ala	Gln	Ile	Arg	
15					635					640					645	
	CCT	GCT	CCT	GAC	GAG	CCT	GTT	CAG	ATG	GAG	GTG	GTT	CAG	GAG	GGG	509
			Pro													
20					650					655					660	
	CCT	GCT	CAG	AAG	GAG	CTG	CTG	CCT	ccc	GTG	GAG	CCT	GCT	CAG	ATG	554
			Gln													
					665					670					675	
25	GTG	GGT	GCC	CAA	ATT	GTA	CTT	GCT	CAC	ATG	GAG	CTG	CCT	CCT	ccc	599
	Val	Gly	Ala	Gln	Ile	Val	Leu	Ala	His	Met	Glu	Leu	Pro	Pro	Pro	
					680					685					690	
			ACT													644
30	Met	Glu	Thr	Ala		Thr	Glu	Val	Ala	Gln	Met	Gly	Pro	Ala	Pro	
					695					700					705	
	ATG	GAA	CCT	GCT	CAG	ATG	GAG	GTT	GCC	CAG	GTA	GAA	TCT	GCT	ccc	689
	Met	Glu	Pro	Ala	Gln	Met	Glu	Val	Ala	Gln	Val	Glu	Ser	Ala	Pro	
35					710					715					720	
			GTG													734
	Met	Gln	Va ₁	Val		Lys	Glu	Pro	Val	Gln	Met	Glu	Leu	Ser	Pro	
4 0			•		725					730					735	

5 CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTT Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Let 755 760 765 TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG 10 Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu 770 775 786 TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg																	
### Pro Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Se		CCC	ATC	GA	G GT	GT	CAG	AAC	GAG	cc:	GT	r cac	ATA	A GA	G CTO	3 TC	T 779
740 745 755 75 5 CCT CCC ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CT Pro Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Let 760 760 760 760 760 760 760 760 760 760		Pro	'Met	Gli	ı Va	l Val	Glr	Lys	Gli	Pro	Va:	l Glr	ı Ile	e Glu	ı .Leı	ı Se	- /// -
### Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Let																75	
### Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Let	_																
### Pro Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Let	5	CCT	cco	ATC.	GAC	GTO	GTC	CAG	AAG	GAZ	cen	GTI	· AAG	ATA	A GAG	CT	3 824
TCT CCT CCC ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG Ser Pro Pro Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu 770 775 775 780 TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 790 790 GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile 800 805 805 TCC AAA AAG CCT CCT CTC CCA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu Ser TTA AGA GAA GAG GCA TCA CCA CCA TCA CCA CTG CCA AAG GAA Ser TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA ASn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Lasn Thr 875 880 885		Pro	Pro	Met	Glu	ı Val	Val	Gln	Lys	Glu	Pro	Val	Lys	Ile	e Glu	Lei	1
Ser Pro Pro Fie Glu Val Val Gln Lys Glu Pro Val Gln Met Glu						755										76	
Ser Pro Pro Fie Glu Val Val Gln Lys Glu Pro Val Gln Met Glu		man															
TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 785 790 790 795 GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile 800 805 810 TCC AAA AAG CCT CCT CTC CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 367 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880	10	TCT	CC1	CCC	ATA	GAG	GTG	GTC	CAG	AAG	GAG	CCT	GTT	CAG	ATG	GAC	869
TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 785 790 790 790 GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT Glu Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile 800 805 810 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 30 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 35 860 865 867 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr	10	ser	Pro	Pro) lle			Val	Gln	Lys	Glu	Pro	Val	Gln	Met	Glı	1
Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 795		,				770					775				•	780)
Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 795		TTG	TOT			3.00	000	-									
GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile 800 805 810 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 660 660 855 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 860 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Lasn Thr 875 885		Leu	Ser	Pro	PEO	Mot	C1	GIG	GTT	CAG	AAG	GAG	CCT	GCT	CAG	AGG	914
GAG CCA CCT CCT CCC AGA GAG CCT CCC CTT CAC ATG GAG CCA ATT Glu Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro Ile 800 805 810 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA AGC GGV Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Lasn Thr 875 880	15			110	710		GIY	vai	Val	Gin			Pro	Ala	Gln	Arg	Ī
Silv Pro Pro Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro 11e 800 805 805 810 810 810 820 805 810 810 810 820 820 820 825						.05					790					795	
Silv Pro Pro Pro Pro Pro Pro Arg Glu Pro Pro Leu His Met Glu Pro 11e 800 805 805 810 810 810 820 805 810 810 810 820 820 820 825		GAG	CCA	CCT	CCT	CCC	AGA	GAG	CCT	ccc		63.6					
20 TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 825 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA AGC Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA SFT Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 875		Glu	Pro	Pro	Pro	Pro	Ara	Glu	Pro	Pro	Tan	CAC	ATG	GAG	CCA	ATT	959
TCC AAA AAG CCT CCT CTC CGA AAA GAT AAA AAG GAA AAG TCT AAC Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 815 820 825 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA Leu Leu Lys Glu Ser Val 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 867 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880							3			110		nis	met	GIU	Pro		
Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 825	20										003					810	
Ser Lys Lys Pro Pro Leu Arg Lys Asp Lys Lys Glu Lys Ser Asn 825		TCC	AAA	AAG	CCT	CCT	CTC	CGA	AAA	GAT	AAA	AAG	ZAD.	DA4	ጥርም	7 7 C	1004
25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830		Ser	Ľys	Lys	Pro	Pro	Leu	Arg	Lys	Asp	Lys	Lvs	Glu	Lvs	Ser	AAC	1004
25 ATG CAG AGT GAA AGG GCA CGG AAG GAG CAA GTC CTT ATT GAA GTT Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 845 850 850 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875										-				-,-	001		
Met Gln Ser Glu Arg Ala Arg Lys Glu Gln Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA GGV Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880																	
Met Gin Ser Glu Arg Ala Arg Lys Glu Gin Val Leu Ile Glu Val 830 835 840 GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880		ATG	CAG	AGT	GAA	AGG	GCA	CGG	AAG	GAG	CAA	GTC	CTT	ATT	GAA	GTT	1049
GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 30 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845		Met	Gln	Ser	Glu	Arg	Ala	Arg	Lys	Glu	Gln	Val	Leu	Ile	Glu	Val	
GGC TTA GTG CCT GTT AAA GAT AGC TGG CTT CTA AAG GAA AGT GTA 30 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845																	
30 Gly Leu Val Pro Val Lys Asp Ser Trp Leu Leu Lys Glu Ser Val 845 850 855 AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885																	
AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885	ın	GGC	TTA	GTG	CCT	GTT	AAA	GAT	AGC	TGG	CTT	CTA	AAG	GAA	AGT	GTA	1094
AGC ACA GAG GAT CTC TCA CCA CCA TCA CCA CCA CTG CCA AAG GAA Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885	,0	GIY	ren	Val	Pro	Val	Lys	Asp	Ser	Trp	Leu	Leu	Lys	Glu	Ser	Val	
Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885						845					850					855	
Ser Thr Glu Asp Leu Ser Pro Pro Ser Pro Pro Leu Pro Lys Glu 860 865 870 AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885		AGC	מרמ	CAC	C N TT	omo.	max										
AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885		Ser '	Thr	GAG	Dan	Crc	TCA	CCA	CCA	TCA	CCA -	CCA	CTG	CCA	AAG	GAA	1139
AAT TTA AGA GAA GAG GCA TCA GGA GAC CAA AAA TTA CTC AAC ACA Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885	5	201		GIU	Asp		ser	PIO	Pro			Pro	Leu	Pro	Lys	Glu	
Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885	-					300					₽ 6 5					870	
Asn Leu Arg Glu Glu Ala Ser Gly Asp Gln Lys Leu Leu Asn Thr 875 880 885		AAT '	TTA	AGA	GAA	GAG	GC»	ጥር አ	GCN -	רא כ	C2 5						
¹ 875 880 885		Asn 1	Leu	Ara	Glu	Glu	Ala	Ser	Glu	Der :	CAA	AAA '	TTA	CTC .	AAC .	ACA	1184
885				- 2					y '			nys .	ren ;	Leu .			
	0										000				;	885	

	GGT	GAA	GGA	AAT	AAA	GAA	GCC	CCT	CTT	CAG	AAA.	GTA	GGA	GCA	GAA	1229
	Gly	Glu	Gly	-Asn	Lys	Glu	Ala	Pro	Leu	Gln	Lys	Val	Gly	Ala	Glu	
					890					895			_		900	
:5	GAG	GCA	GAT	GAG	AGC	CTA	CCT	GGT	CTT	GCT	GCT	AAT	ATC	.AAC	GAA	1274
					Ser											
					905					910					915	
	TCT	ACC	CAT	ATT	TCA	TCC	TCT	GGA	CAA	AAC	TTG	AAT	ACG	CCA	GAG	1319
10					Ser											
					920					925			: -		930	
	GGT	GAA	ACT	TTA	AAT	GGT	AAA	CAT	CAG	ACT	GAC	AGT	ATA	GTT	TGT	1364
					Asn											
15					935					940					945	
	GAA	ATG	AAA	ATG	GAC	ACT	GAT	CAG	AAC	ACA	AGA	GAG	AAT	CTC	ACT	1409
	Glu	Met	Lys	Met	Asp	Thr	Asp	Gln	Asn	Thr	Arg	Glu	Asn	Leu	Thr	
					950					955					960	
20																
	GGT	ATA	AAT	TCA	ACA	GTT	GAA	GAA	CCA	GTT	TCA	CCA	ATG	CTT	CCC	1454
	Gly	Ile	Asn	Ser	Thr	Val	Glu	Glu	Pro	Val	Ser	Pro	Met	Leu	Pro	
					965					970					975	
25	CCT	TCA	GCA	GTA	GAA	GAA	CGT	GAA	GCA	GTG	TCC	AAA	ACT	GCA	CTG	1499
	Pro	Ser	Ala	Val	Glu	Glu	Arg	Glu	Ala	Val	Ser	Lys	Thr	Ala	Leu	
					980					985					990	
																1544
30	Ala	Ser	Pro	Pro	Ala	Thr	Met	Ala				Ser	Gln	Glu	Ile	
					995					1000)				1005	i
							•									
																1589
	Asp	Glu	Asp	Glu	Gly	Ile	His	Ser	His	Glu	Gly	Ser	Asp	Leu	Ser	
35					1010)				1015	5				1020)
																1634
	Asp	Asn	Met,	Ser	Glu		Ser	Asp	Asp	Ser	Gly	Leu	His	Gly	Ala	
40			•		1025	5				1030)				1035	i
40																

	CGG	CCA	GTT	CCA	CAA	GAA	TCT	AGC	AGA	AAA	AAT	GÇA	AAG	GAA	GCC	1679
	Arg	Pro	Val	Pro	Gln	Glu	Ser	Ser	Arg	Lys	Asn	Ala	Lys	Glu	Ala	
					.104					104			_		105	0
.5	TTG	GCA	GTC	AAA	GCG	GCT	AAG	GGA	GAT	TTT	GTT	TGT	ATC	TTC	TGT	1724
	Leu	Ala	Val	Lys	Ala	Ala	Lys	Gly	Asp	Phe	Val	Cys	Ile	Phe	Cys	
					1055					106					106	5
	GAT	CGT	TCT	TTC	AGA	AAG	GGA	AAA	GAT	TAC	AGC	AAA	CAC	CTC	AAT	1769
10	Asp	Arg	Ser	Phe	Arg	Lys	Gly	Lys	Asp	Tyr	Ser	Lys	His	Leu	Asn	
					1070					1075					1080)
	CGC	CAT	TTG	GTT	TAA	GTG	TAC	TAT	CTT	GAA	GAA	GCA	GCT	CAA	GGG	1814
	Arg	His	Leu	Val	Asn	Val	Tyr	Tyr	Leu	Glu	Glu	Ala	Ala	Gln	Gly	
15					1085					1090					1095	
	CAG	GAG	TAAT	rg aa	ACTI	TGAA	CAA	GGTI	TCA	GTTC	TTAG	TT				1855
	Gln	Glu														
20																
	TGTA	AGGI	'AT A	TTAC	ATTT	T AI	ATTO	ATTI	ATG	ATAG	CAG	ACAA	CCTT	TT		1905
	AAGA	TTGC	TT T	'AATT	'AGTA	T CT	GATG	TTGA	TTT	TTAA	GTG	GCAT	TCTT	TT		1955
25	CCTT	AGGA	CT T	TTTA	TGTA	T AC	CTGT	TGAT	TGT	TGTG	TAA	ATTT	TAGT	AA		2005
																•
	ATCT	AAGA	GA G	TGTA	CTAA	A CC	AGCA	GGTA	TCT	GTTA	GCT	TATG'	TGTT	TA		2055
	ATTG	TAAA	TA G	AAGG	CTAA	G AT	GGTA	TAAC	AGC.	ATTT	TAT	TGCT	TTGT	CC		2105
30																
	AGCT.	ACAA	CA T	GTCA	TTTT	T TT	CTCC	ATGT	CTT	ATCT	TCC '	TGTT:	CAC	TT	:	2155
	TAGT	TAT	TC T	TCGT	TTTT	T AT	TGAG.	ATCT	ATA	AAAA	ATT (GCT.	ract"	TA	:	2205
35	ATAG	CAAA	TT A	CTTG.	AAGA	A TT	TGCC'	TGCT	TTA	TATA:	AAG 1	TTAG	CACT	ΓT	:	2255
							-									
	AAGA'	TTTT	TT T	TTTA	GAGA:	r ga	GAAG	ACAT	TTA	TTAP	SAA (SAAAS	ATT	CC	2	2305
				į												
	CCCA	GCAA!	TA G	ACAG'	TCTA:	CA(GTCC	AAGT	ATT	ract:	rcc :	rgagi	TTT	GA	2	2355
40																
	TCAA	TATT:	TT T	TATT:	TGTG:	T AT	GTTA	ATCG	TCAT	LAAA 1	AC A	AGTGA	TTT	rg	2	405

	GTGTGTTTTT	TATTTTGGTG	CTTTAATGGC	TTAAGATGTT	GCACATTTTT	.2455
	TTTTTCTTTT	GGTTTCTGTT	TATGTTTTT	TGCCTATGCA	GTTAAATTTT	2505
5	TCCTAGAAAT	AGCATTTGTG	TTGAACAGTA	ACACTTTATA	CATATATATA	2555
	TGCATGTTTA	TTTTGTTTGG	CGTCTTTGGA	GGGATGCTTT	TAGACTTGTT	2605
10	TGCAAAAGGG	CAGTTTTCTT	TTTCTTTGCT	GCAGTTGTCT	ATTTTGCAGA	2655
	ATAATAGTGT	GTGCAAGTTT	GTGAGCAAAT	GAAATATGCA	GGTTCAATCT	2705
	ATTGATTTTG	ATTTTTACAT	СТТАТАТСТА	TGCCAGAATC	TGTATTTCAT	2755
15	ATAACTTATT	TATTTCGAAT	GGATGTAGTA	AATTCACAGC	TATCAGTTTT	2805
	GATTTTGCAA	TAAATAAACC	ACTAGGTTGC	ATGTCGAACA	AATTTTTATC	2855
20	TCAAATACCA	ACCATCAGTT	TTTTTTTCA	TGTGTTTTGG	TACAGCTAAT	2905
	TCCTAATTGT	AGAGTGTTAA	ATGTTTGAGG	AGAACCTTTT	CTCATAGATG	2955
	GTTGGTGTTC	ATATGGCNAC	TTTACAATAA	AGAGAACTGT	AAGTGATATT	3005
25	TGGAAACTAC	AAACCTGGAA	TTAGGAGATA	TAATTATTCC	TTCAAGTTTT	3055
	ATAGATATCA	CTTGGGAGAT	TCCAAAGCCA	TAGCTATTAC	GCNGCAAACC	3105
30	TAGGATAAGA	AAGGTAGTAT	GAGTGCTGGT	AGACCAGCTG	CAACATTTCC	3155
	TATATCAGAT	GAAAAAGGCT	GGTGAAACAA	GTACAGTCCA	GATTTTTTAA	3205
	AATCATACTT	TCTCAGGGAT	CTCCACAAAC	TGGTGGGTGT	CCTGGCTGTC	3255
35	TGTGTGATAG	CCTCTTTCTA	TAGGTGAGGC	CTCAAATGAA	TTGCAGCTAT	3305
	CCTGGTGTTC	CTATGAGGGC	ACTTGTATGA	AAAAGGCAGT	ACTCCAAAAC	3355
40	ATTTTTGATG	GTTCTTTGGC	CAGTTGCCAA	AGAGTGTGAA	AGAATCCAAT	3405
70	AGAGGATTTT	TCTTACTGAT	AGCAGTCATT	CATTGCAGTA	ATAAAATA	3455

	- 03 -	
	TGAATTCCCA TTAGGGAATC TTGAATTCTG ACCTCCCATA CTCCGTTTTG	.3505
	AAATAACCAC TTATATTTCA TTTTTTAAAA ATCTGATGAT CTCTTTGAGG	3555
·5	CAGGTTTCAG ATTTGGCAGT ACAACATGAA AGATTAGGAA AAGCATTAAT	3605
	AACGTGTGGG TGGAAAGCTT GTTAAAAATC TGAGAGTGAA GTTTGAGTTA	3655
10	AAAGTTGTTT GACATGGCAT TGACTGGGAG GCCAAAGATT TAAAGAAGCG	3705
10		
	(2) INFORMATION FOR SEQ ID NO: 8:	
	(i) SEQUENCE CHARACTERISTICS	
	(A) LENGTH: 24 base pairs	
	(B) TYPE: nucleic acid	
15	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: Other nucleic acid	
	(iii) HYPOTHETICAL: no	
	(iv) ANTI-SENSE: no	
20	(x) PUBLICATION INFORMATION:	
	(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toled	>-
	Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler	,
	Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Ga	il
	(B) TITLE: REST: A Mammalian Silencer Protein that Restric	cts
25	Sodium Channel Gene Expression to Neurons	
	(C) JOURNAL: Cell	
	(D) VOLUME: 80	
	(E) ISSUE:	
20	(F) PAGES:	
30	(G) DATE: March 24, 1995	
	(K) RELEVANT RESIDUES IN SEQ ID NO:8:FROM 1 TO 24	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
35	TGYAARCCNT GYCARTAYGA RGCN	24
رر	(2) INFORMATION FOR CRO TO	
	(2) INFORMATION FOR SEQ ID NO: 9:	
	(i) SEQUENCE CHARACTERISTICS	
	(A) LENGTH; 24 base pairs	
40	(B) TYPE: hucleic acid	
75	(C) STRANDEDNESS: single	

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid --

- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: no
- (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-5 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 10 (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:9:FROM 1 TO 24
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

NGTYTTRTAR TCRCARTGNG GRCA

- (2) INFORMATION FOR SEQ ID NO: 10:
- 20 (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 3291 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
- 30 (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-
- 35 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 40 (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:

(G) DATE: March .24, .1995																
	(K) RELEVANT RESIDUES IN SEQ ID NO:10:FROM 1 TO 3291 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:															
	(xi) SE	QUE	ICE I	ESCR	IPTI	ON:	SEQ	ID N	10:10):					
5	ATG	GCC	ACC	CAG	GTA	ATG	GGG	CAG	TCI	TCI	GGA	GGA	GGA	GGG	CTG	45
	Met	Ala	Thr	Gln	Val	Met	Gly	Gln	Ser	Ser	Gly	Gly	Gly	Gly	Leu	
	1				5					10					15	
10	TTT	ACC	AGC	AGT	GGC	AAC	ATT	GGA	ATG	GCC	CTG	CCT	AAC	GAC	ATG	90
10	Ρ.	he T	hr S	er S		ly A	sn I	le G	ly M	let A	la L	eu P	TO A	sn A	sp Me	≥ t
					20					25				-	30	
	m s m	~~~														
	TAL	GAC	TIG	CAT	GAC	CIT	TCC	AAA	GCT	GAA	CTG	GCC	GCA	CCT	CAG	135
. 15	Tyr	АБР	Leu	His		Leu	Ser	Lys	Ala	Glu	Leu	Ala	Ala	Pro	Gln	
					35					40					45	
	CTT	ידידע	ATG	CTC	CCX	7 7 TO	C.T.C									
	Leu	Ile	Met	CTG	Ala	AAI	U-1	GCC NI-	TTA	ACT	GGG	GAA	GTA	AAT	GGC	180
				Leu	50	ASII	VAI	Ala	Leu		GIA	Glu	Val	Asn		
20					50					55					60	
	AGC	TGC	TGT	GAT	TAC	CTG	GTC	CCT	CAN	CAA	202	C N C	3 mc	222		
	Ser	Cys	Cys	Asp	Tvr	Leu	Val	Glv	GAA	GAA	AGA	CAG	ATG	GCA	GAA	225
		_	-	•	65			Cly	, J	70	Arg	GIII	mec	AIA		
										. 0					75	
25	CTG	ATG	CCG	GTT	GGG	GAT	AAC	AAC	TTT	TCA	GAT	AGT	CDD	ממ	GGN	270
	Leu	Met	Pro	Val	Gly	Asp	Asn	Asn	Phe	Ser	Asp	Ser	Glu	Glu	Glv	270
					80					85					90	
•	GAA	GGA	CTT	GAA	GAG	TCT	GCT	GAT	ATA	AAA	GGT	GAA	CCT	CAT	GGA	315
30				Glu												
					95					100					105	
	CTG	GAA	AAC	ATG	GAA	CTG	AGA	AGT	TTG	GAA	CTC	AGC	GTC	GTA	GAA	360
	Leu	Glu	Asn	Met	Glu	Leu	Arg	Ser	Leu	Glu	Leu	Ser	Val	Val	Glu	
35					110					115					120	
	CCT															405
	Pro	Gln	Pro	Vaļ	Phe	Glu	Ala	Ser	Gly	Ala	Pro	Asp	Ile	Tyr	Ser	

130

135

	TCA	דממ.	מממ	CCT	Color	CCC	COT	C N N	202	~~~		-				
															AAA	450
	361	ASII	Lys	ATA		ALA	PTO	GIU	Thr		GIA	Ala	Glu	Asp	_	
					140					145					150	
:5	GGC	AAG	.AGC	TCG	AAG	ACC	AAA	ccc	TTT	CGC	TGT	AAG	CCA	TGC	CAA	495
	Gly	Lys	Ser	Ser	Lys	Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	
					155					160					165	
	TAT	GAA	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC	ATC	AGA	GTT	540
10	Tyr	Glu	Ala	Glu	Ser	Glu	Glu	Gln	Phe	Val	His	His	Ile	Arg	Val	
					170					175					180	
			GCT													585
. .	His	Ser	Ala	Lys	Lys	Phe	Phe	Val	Glu	Glu	Ser	Ala	Glu	Lys	Gln	
15					185					190					195	
	GCA	AAA	GCC	AGG	GAA	TCT	GGC	TCT	TCC	ACT	GCA	GAA	GAG	GGA	GAT	630
	Ala	Lys	Ala	Arg	Glu	Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
20					200					205				_	210	
	TTC	TCC	AAG	GGC	CCC	ATT	CGC	TGT	GAC	רפר	ፕሮር	GGC	ጥ ልር	דממ	ACT	675
			Lys													0,5
			•	-	215		3	-7-	F	220	-75	U	- , -	7.011	225	
25	AAT	CGA	TAT	GAT	CAC	TAT	ACA	GCA	CAC	CTG	AAA	CAC	CAC	ACC	AGA	720
	Asn	Arg	Tyr	Asp	His	Tyr	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	
					230					235					240	
	GCT	GGG	GAT	AAT	GAG	CGA	GTC	TAC	AAG	TGT	ATC	ATT	TGC	ACA	TAC	765
30	Ala	Gly	Asp	Asn	Glu	Arg	Val	Tyr	Lys	Cys	Ile	Ile	Cys	Thr	Tyr	
					245					250					255	
	ACA	ACA	GTG	AGC	GAG	TAT	CAC	TGG	AGG	AAA	CAT	TTA	AGA	AAC	CAT	810
	Thr	Thr	Val	Ser	Glu	Tyr	His	Trp	Arg	Lys	His	Leu	Arg	Asn	His	
35					260					265					270	
	TTT	CCA	AGG	AAA	GTA	TAC	ACA	TGT	GGA	AAA	TGC	AAC	TAT	TTT	TCA	855
	Phe	Pro	Arg	Lys	Val	Tyr	Thr	Cys	Gly	Lys	Cys	Asn	Tyr	Phe	Ser	
40					275					280					285	
mil.																

	GA	.AG	AA.	AA. A	CAA.	LAT T	GT.	CAC	CAT	r gr	T AG	A AC	r car	r aca	A GG	A 900
	Asp	Arg	Ly	s Ası	Ası	туг	Va]	l Glr	His	s Va	l Ar	Thi	r His	Thi	Gly	,
					290					.29					300	
5		CGC	: cci	A TAI	AAA 1	TGI	GAZ	CTI	TGI	r cc	TAC	TC	A AGT	TCI	CAG	945
	Glu	Arg	Pro	Tyr	Lys	Cys	Glu	ı Leu	Cys	Pro	тут	Ser	Ser	Ser	Glr	1
					305					310					315	
	AAG	ACT	CAI	CTA	ACT	AGA	CAI	ATG	CGT	' ACI	CAT	TCA	GGT	GAG	AAG	990
10	Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thr	His	Ser	Gly	Glu	Lys	
				•	320					325			_	•	330	
	CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	TAA	CAA	CAT	1035
	Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
15					335					340					345	
	GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	1080
	Glu	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lys	Pro	
30					350					355					360	
20																
	CTT -	AAT	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	1125
	Leu	Asn	Cys	Pro		Cys	Asp	Tyr	Lys	Thr	Ala	Asp	Arg	Ser	Asn	
					365					370					375	
25	mmo				_											
23	TTC	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	TAA	1170
	Pne	rys	rys	His		Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
					380					385					390	
	TCC	CCT	Cm.	mam												
30	Cyc	Pro	UAL	TGT	GAC	TAT	GCA	GCT	TCC	AAG	AAG	TGT	AAT	CTA	CAG	1215
20	Cys	PIO	vai	Cys		ıyr	Ala	Ala	Ser		Lys	Cys	Asn	Leu	Gln	
					395					400					405	
	ፐልጥ	CAC	тт С	***	330	220	01 m									
	Tur	Hic	Pho	AAA	101	AAG	CAT	CCT	ACT	TGT	CCT	AAT	AAA	ACA	ATG	1260
35	-71	*****	FIIE	Lys		Lys	HIS	Pro			Pro	Asn	Lys	Thr	Met	
					410					415					420	
	ርልጥ	CTC	י. י	ממת	CTC	7 N N										
	Asn	บลา	20×	AAA	375 J	AAA T	CIA	AAG .	AAA	ACC	AAA	AAA -	CGA	GAG	GCT	1305
		- 41	Jer	Lys	vai 425	nys	nen	ьys			rys	Lys	Arg	Glu .	Ala	
40					763					430					435	

	GAC	TTG	CCT	GAT	AAT	ATT	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	.1350
	Asp	Leu	Pro	Авр	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
					440					445					450	
5	ACA	AAA;	ATA	AAA	GGG	GAT	GTG	GCT	GGA	AAG	AAA	AAT	GAA	AAG	TCC	.1395
	Thr	Lys	Tle	Lys	Gly	Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
					455					460					-465	
																1440
10	Val	Lys	Ala	Glu	Lys	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
					470					475			•	·	480	
																1485
15	Asn	Asn	Val	Ser		Ile	Gln	Val	Thr	Thr	Arg	Thr	Arg	Lys	Ser	
15					485					490					495	
																1530
	vai	Thr	GIU	vai		Glu	Met	Asp	Val		Thr	Gly	Ser	Asn		
20					500					505					510	
20	CDD	מממ	الململة	እርጥ	מממ	א כיידי	770	222	200		300		ama	~~~	~~~	
							Lys									1575
	OI u	Lys	FIIC	361	515	1111	Lys	гуs	ser	520	Arg	ьуs	Leu	GIU		
					313					320					525	
25	GAC	AGC	CAT	TCT	TTA	CAT	GGT	CCT	GTG	דבב	CAT	GAG	CAD	тст	ጥርል	1620
							Gly									1020
	•				530		,			535		0	014	001	540	
	ACA	AAA	AAG	AAA	AAG	AAG	GTA	GAA	AGC	AAA	TCC	AAA	AAT	AAT	AGT	1665
30							Val									
					545										555	
	CAG	GAA	GTG	CCA	AAG	GGT	GAC	AGC	AAA	GTG	GAG	GAG	AAT	AAA	AAG	1710
	Gln	Glu	Val	Pro	Lys	Gly	Asp	Ser	Lys	Val	Glu	Glu	Asn	Lys	Lys	
35					560					565					570	
	CAA	AAT	ACT	TGC	ATG	AAA	AAA	AGT	ACA	AAG	AAG	AAA	ACT	CTG	AAA	1755
	Gln	Asn	Thr	Cys	Met	Lys	Lys	Ser	Thr	Lys	Lys	Lys	Thr	Leu	Lys	
			·		575					580					585	
40																

	AA:	r .aa.	A TC	A AG	r aac	AAA E	A AGO	AG:	AA 1	G CC	T CC	CAC	AAC	GA	A CC	T .1800
	Ası	Ly	s Se	r Se	r Lys	Lys	Ser	Ser	Lys	s Pr	o Pro	Glr	Lys	s Glı	ı Pro	5
					590					.59			•		600	
															50,	•
:5	GTT	GA	AA E	G GGZ	TCI	GCI	CAG	ATO	GAC	c	ר ככי	CAG	ATC	: פכני	י רריז	1845
	Val	Glu	ı Lys	s Gly	/ Sei	Ala	Gln	Met	Ast	Pro	Pro	Glr	Met	G)	Dro	
					605	5			•	610					615	
															011	•
	GCT	ccc	ACA	GAG	GCG	GTT	CAG	AAG	GGG	3 CCC	GTT	. CAG	GTG	G A G	CTC	1890
10	Ala	Pro	Thr	Glu	Ala	Val	Gln	Lys	Gly	Pro	Val	Gln	Val	Glu	T.e.	1030
					620			_	•	625			,,,	GIU	630	
	•														630	
	CCA	CCI	. ccc	ATG	GAG	CAT	GCT	CAG	ATG	GAG	GGT	GCC	CAG	מדמ	ccc	1935
	Pro	Pro	Pro	Met	Glu	His	Ala	Gln	Met	Glu	Glv	Ala	Gln	Tle	750	1935
15					635					640			0111	116	645	
															045	
	CCT	GCT	CCI	GAC	GAG	CCT	GTT	CAG	ATG	GAG	GTG	ىلىدى	CAG	GNG	ccc	1980
	Pro	Ala	Pro	Asp	Glu	Pro	Val	Gln	Met	Glu	Val	Val	Gln	GAG	Glar	1380
					650					655		• • • •	0211	GIU	660	
20															880	
	CCT	GCT	CAG	AAG	GAG	CTG	CTG	CCT	ccc	GTG	GAG	CCT	GCT	CAG	N TC	2025
	Pro	Ala	Gln	Lys	Glu	Leu	Leu	Pro	Pro	Val	Glu	Pro	Ala	Gln	Met	2023
					665					670			,,,,	0111	675	
															0,75	
25	GTG	GGT	GCC	CAA	ATT	GTA	CTT	GCT	CAC	ATG	GAG	CTG	ССТ	ССТ	CCC	2070
	Val	Gly	Ala	Gln	Ile	Val	Leu	Ala	His	Met	Glu	Leu	Pro	Pro	Pro	2070
					680					685					690	
															0,50	
	ATG	GAG	ACT	GCT	CAG	ACG	GAG	GTT	GCC	CAA	ATG	GGG	ССТ	GCT	רכר	2115
30	Met	Glu	Thr	Ala	Gln	Thr	Glu	Val	Ala	Gln	Met	Glv	Pro	Ala	Pro	
					695					700		2			705	
															, 05	
	ATG	GAA	CCT	GCT	CAG	ATG	GAG	GTT	GCC	CAG	GTA	GAA	тст	CCT	CCC	2160
	Met	Glu	Pro	Ala	Gln	Met	Glu	Val	Ala	Gln	Val	Glu	Ser	Ala	Pro	2100
35					710					715					720	
															. 20	
	ATG	CAG	GTG	GTC	CAG	AAG	GAG	CCT	GTT	CAG	ATG	GAG (CTG '	ተርተ	ССТ	2205
	Met	Gln	Val	Val	Gln	Lys	Glu	Pro	Val	Gln	Met	Glu	Len	Ser	Pro	4403
				1	725					730	-				735	
40																

	CCC	ATG	GAG	GTG	GTC	CAG	AAG	GAG	CCT	GTT	CAG	ATA	GAG	CTG	TCT	.225
											Gln					
					740					745					750	
:5	CCT	CCC	ATG	GAG	GTG	GTC	CAG	AAG	GAA	CCT	GTT	AAG	ATA	GAG	CTG	:2299
	Pro	Pro	Met	Glu	Val	Val	Gln	Lys	Glu	Pro	Val	Lys	Ile	Glu	Leu	
					755					760					765	
	TCT	CCT	CCC	ATA	GAG	GTG	GTC	CAG	AAG	GAG	CCT	GTT	CAG	ATG	GAG	2340
10	Ser	Pro	Pro	Ile	Glu	Val	Val	Gln	Lys	Glu	Pro	Val	Gln	Met	Glu	
					770					775					780	
	TTG	TCT	CCT	CCC	ATG	GGG	GTG	GTT	CAG	AAG	GAG	CCT	GCT	CAG	AGG	2385
15	Leu	Ser	Pro	Pro		Gly	Val	Val	Gln	Lys	Glu	Pro	Ala	Gln	Arg	
15					785					790					795	
	63.6	~~~														
											CAC					2430
	GIU	PIO	PIO	Pro		Arg	Glu	Pro	Pro	Leu	His	Met	Glu	Pro	Ile	
20					800					805					810	
20	TCC	מממ	ስ ክ'C	CCT		~~~	001					<u>a</u>				
											AAG					2475
	561	Dys	Lys	PIO	815	ren	Arg	Lys	Asp		Lys	Glu	Lys	Ser		
					613					820					825	
25	ATG	CAG	AGT	GAA	AGG	GCA	ccc	אמ	GNG	רח ח	GTC	C-CT-CT-CT-CT-CT-CT-CT-CT-CT-CT-CT-CT-CT	n mm	<i>-</i>		2520
											Val					2520
					830			2,5	014	835	Val	Deu	TIE	GIU	840	
										درن					5 4 U	
	GGC	TTA	GTG	CCT	GTT	AAA	GAT	AGC	TGG	СТТ	CTA	DAA	AAD	AGT	СТЪ	2565
30											Leu					2303
	_				845		•			850		-1-			855	
	AGC	ACA	GAG	GAT	CTC	TCA	CCA	CCA	TCA	CCA	CCA	CTG	CCA	AAG	GAA	2610
											Pro					
35					860					865				•	870	
															_	
	TAA	TTA	AGA	GAA	GAG	GCA	TCA	GGA	GAC	CAA.	AAA	TTA	CTC	AAC	ACA	2655
											Lys					
					875					880					885	
10														•		

	GGI	GAZ	A GGZ	(AA:	AAA 1	GAA	GCC	CC1	CI	CAC	AAA	GTA	GGA	GC	GAA	:2700
	Gly	' Glı	ı Gly	' Asi	l Lys	Glu	Ala	Pro	Let	ı Glm	Lys	Val	Gly	Ala	Glu	l
					890					895					900	
.5	GAG	GCZ	GAT	GAC	AGC	CTA	CCI	GGI	. CII	GCI	GCI	TAA	ATO	·AAC	GAA	.2745
	Glu	Ala	·Asp	Glu	Ser	Leu	Pro	Gly	Let	Ala	Ala	Asn	Ile	Asn	Glu	
					905					910					915	
	TCT	ACC	CAT	ATI	TCA	TCC	TCT	GGA	CAA	AAC	TTG	AAT	ACG	CCA	GAG	2790
10	Ser	Thr	His	Ile	Ser	Ser	Ser	Gly	Gln	Asn	Leu	Asn	Thr	Pro	Glu	2,50
					920			•		925					930	
															230	
	GGT	GAA	ACT	TTA	AAT	GGT	AAA	CAT	CAG	ACT	GAC	AGT	ATA	GTT	TGT	2835
	Gly	Glu	Thr	Leu	Asn	Gly	Lys	His	Gln	Thr	Asp	Ser	Tle	Val	Cve	2033
15					935					940				• • • •	945	
															343	
	GAA	ATG	AAA	ATG	GAC	ACT	GAT	CAG	AAC	ACA	AGA	GAG	AAT	CTC	אריד	2880
	Glu	Met	Lys	Met	Asp	Thr	Asp	Gln	Asn	Thr	Arg	Glu	Asn	Len	Thr	2000
					950		_			955	5			200	960	
20						•									300	
	GGT	ATA	AAT	TCA	ACA	GTT	GAA	GAA	CCA	GTT	TCA	CCA	ATG	ىلىك	CCC	2925
	Gly	Ile	Asn	Ser	Thr	Val	Glu	Glu	Pro	Val	Ser	Pro	Met	Len	Pro	2925
					965					970				200	.975	
										•					.915	
25	CCT	TCA	GCA	GTA	GAA	GAA	CGT	GAA	GCA	GTG	TCC	AAA	ACT	GCA	CTG	2970
	Pro	Ser	Ala	Val	Glu	Glu	Arg	Glu	Ala	Val	Ser	Lvs	Thr	Δla	Len	23,0
					980					985		-,,-			990	
															230	
	GCA	TCA	CCT	CCT	GCT	ACA	ATG	GCA	GCA	AAT	GAG	тст	CAG	GAA	ידינ	3015
30	Ala	Ser	Pro	Pro	Ala	Thr	Met	Ala	Ala	Asn	Glu	Ser	Gln	Glu	Tle	3013
					995					1000					1005	
															1003	
	GAT	GAA	GAT	GAA	GGC	ATC	CAC	AGC	CAT	GAA	GGA	AGT	GAC	СТЪ	АСТ	3060
	Asp	Glu	Asp	Glu	Gly	Ile	His	Ser	His	Glu	Glv	Ser	ASD	Leu	Ser	3000
35					1010					1015				Dea	1020	
															1020	
	GAC	AAC	ATG	TCA	GAG	GGT	AGT	GAT	GAT	TCT	GGA	TTG	САТ	GGG	ርርጥ	3105
	Asp	Asn	Met	Ser	Glu	Gly	Ser	Asp	Asp	Ser	Glv	Len	His	Glv	Ala	-103
				ì	1025			•		1030					103E	

CGG CCA GTT CCA CAA GAA TCT AGC AGA AAA AAT GCA AAG GAA GCC 3150 Arg Pro Val Pro Gln Glu Ser Ser Arg Lys Asn Ala Lys Glu Ala 1040 .1045 1050

5 TTG GCA GTC AAA GCG GCT AAG GGA GAT TTT GTT TGT ATC TTC TGT 3195 Leu Ala Val Lys Ala Ala Lys Gly Asp Phe Val Cys Ile Phe Cys 1055 1060 1065

GAT CGT TCT TTC AGA AAG GGA AAA GAT TAC AGC AAA CAC CTC AAT 3240

10 Asp Arg Ser Phe Arg Lys Gly Lys Asp Tyr Ser Lys His Leu Asn

1070 1075 1080

CGC CAT TTG GTT AAT GTG TAC TAT CTT GAA GAA GCA GCT CAA GGG 3285 Arg His Leu Val Asn Val Tyr Tyr Leu Glu Glu Ala Ala Gln Gly 15 1085 1090 1095

CAG GAG

Gln Glu

1097

- (2) INFORMATION FOR SEO ID NO: 11:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 63 base pairs
- 25 (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
- 30 (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
- 35 (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- 40 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell

- 73 -

```
(D) VOLUME: 80
```

- (E) ISSUE:
- (F) PAGES:
- (G) DATE: March 24, 1995
- (K) RELEVANT RESIDUES IN SEQ ID NO:11:FROM 1 TO 63
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGT AAG CCA TGC CAA TAT

Cys Lys Pro Cys Gln Tyr

10

165

GAA GCA GAA TCT GAA GAA CAG TTT GTG CAT CAC ATC AGA GTT CAC
Glu Ala Glu Ser Glu Glu Gln Phe Val His His Ile Arg Val His
170 175 180

- (2) INFORMATION FOR SEQ ID NO: 12:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 63 base pairs
 - (B) TYPE: nucleic acid
- 20 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 25 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
- 30 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
- 35 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES: 4
- 40 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:12:FROM 1 TO 63
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:...

- 74 -

TGT GAC CGC TGC GGC TAC AAT ACT
Cys Asp Arg Cys Gly Tyr Asn Thr
220 .225

.24

5 AAT CGA TAT GAT CAC TAT ACA GCA CAC CTG AAA CAC CAC Asn Arg Tyr Asp His Tyr Thr Ala His Leu Lys His His 230 235

63

- (2) INFORMATION FOR SEQ ID NO: 13:
- 10 (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 63 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 15 (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
- 20 (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-25 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 30 (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:13:FROM 1 TO 63
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGT ATC ATT TGC ACA TAC
Cys Ile Ile Cys Thr Tyr
250 255

- 75 -

ACA ACA GTG AGC GAG TAT CAC TGG AGG AAA CAT TTA AGA AAC CAT Thr Thr Val Ser Glu Tyr His Trp Arg Lys His Leu Arg Asn His 260 .265 .270

- (2) INFORMATION FOR SEQ ID NO: 14:
 - (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 63 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
- 10 (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
- 15 (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
- 20 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
- 25 (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
- 30 (K) RELEVANT RESIDUES IN SEQ ID NO:14:FROM 1 TO 63 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGT GGA AAA TGC AAC TAT TTT TCA Cys Gly Lys Cys Asn Tyr Phe Ser 35

24

280 285

GAC AGA AAA AAC AAT TAT GTT CAG CAT GTT AGA ACT CAT 63 Asp Arg Lys Asn Asn Tyr Val Gln His Val Arg Thr His 290 295

- (2) INFORMATION FOR SEQ ID NO: 15:
- (i) SEQUENCE CHARACTERISTICS

(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: no

(ii) MOLECULE TYPE: cDNA to mRNA

	(A) LENGTH: 63 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	
	(D) TOPOLOGY: linear	
:5	(ii) MOLECULE TYPE: cDNA to mRNA	
	(iii) HYPOTHETICAL: no	
	(iv) ANTI-SENSE: no	
	(vi) ORIGINAL SOURCE:	
	(A) ORGANISM: Human	
10	(H) CELL LINE: HeLa	
	(vii) IMMEDIATE SOURCE:	
	(A) LIBRARY: cDNA	
	(x) PUBLICATION INFORMATION:	
	(A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-	-
15	Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,	
	Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail	L
	(B) TITLE: REST: A Mammalian Silencer Protein that Restrict	
	Sodium Channel Gene Expression to Neurons	
	(C) JOURNAL: Cell	
20	(D) VOLUME: 80	
	(E) ISSUE:	
	(F) PAGES:	
	(G) DATE: March 24, 1995	
	(K) RELEVANT RESIDUES IN SEQ ID NO:15:FROM 1 TO 63	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:	
	TGT GAA CTT TGT CCT TAC TCA AGT TCT CAG	30
	Cys Glu Leu Cys Pro Tyr Ser Ser Ser Gln	
	310 315	
30		
	AAG ACT CAT CTA ACT AGA CAT ATG CGT ACT CAT	63
	Lys Thr His Leu Thr Arg His Met Arg Thr His	
	320 325	
35	(2) INFORMATION FOR SEQ ID NO: 16:	
	(i) SEQUENCE CHARACTERISTICS	
	(A) LENGTH: 66 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double	

(H) CELL LINE: HeLa (vii) IMMEDIATE SOURCE: (A) LIBRARY: cDNA

- 77	-
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: Human	
(H) CELL LINE: HeLa	
(vii) IMMEDIATE SOURCE:	
(A) LIBRARY: cDNA	
(x) PUBLICATION INFORMATION:	
(A) AUTHORS: Chong, Jayhong A.,	Tania-Ramirez José molodo
Aral, Juan, Zheng, Yingcong, Boutro	s. Michael C Altechyler
Yelena M., Frohman, Michael A., Kran	ner, Susan D. Mandel Gail
(B) TITLE: REST: A Mammalian Sile	encer Protein that Pestriate
Sodium Channel Gene Expression to Ne	eurons
(C) JOURNAL: Cell	
(D) VOLUME: 80	
(E) ISSUE:	
(F) PAGES:	
(G) DATE: March 24, 1995	
(K) RELEVANT RESIDUES IN SEQ ID N	O:16:FROM 1 TO 66
(xi) SEQUENCE DESCRIPTION: SEQ ID NO):16:
TGT GAT CAG TGC AGT TAT GTG GCC TCT	
Cys Asp Gln Cys Ser Tyr Val Ala Ser	Asn Gln His
335 340	345
Cha cma and con	
GAA GTA ACC CGC CAT GCA AGA CAG GTT	
Glu Val Thr Arg His Ala Arg Gln Val	His
350	355
(2) INFORMATION FOR SEQ ID NO: 17	
(i) SEQUENCE CHARACTERISTICS	:
(A) LENGTH: 63 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: double	
(D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: cDNA to mRNA	
(iii) HYPOTHETICAL: no	
(iv) ANTI-SENSE: no	
(vi) ORIGINAL SOURCE:	
(A) ORGANISM: Human	

- (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- 5 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
- 10 (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:17:FROM 1 TO 63
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
- 15 TGC CCA CAC TGT GAT TAC AAA ACA GCA GAT AGA AGC AAC

 Cys Pro His Cys Asp Tyr Lys Thr Ala Asp Arg Ser Asn

 365 370 375

TTC AAA AAA CAT GTA GAG CTA CAT
20 Phe Lys Lys His Val Glu Leu His
380

- (2) INFORMATION FOR SEQ ID NO: 18:
- (i) SEQUENCE CHARACTERISTICS
- 25 (A) LENGTH: 66 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
- 30 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
- 35 (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
 40 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
 Sodium Channel Gene Expression to Neurons

- (C) JOURNAL: Cell
- (D) VOLUME: 80
- (E) ISSUE:
- (F) PAGES:
- 5 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:18:FROM 1 TO 66
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGC CCT GTA TGT GAC TAT GCA GCT TCC AAG AAG TGT AAT CTA CAG

10 Cys Pro Val Cys Asp Tyr Ala Ala Ser Lys Lys Cys Asn Leu Gln

395 400 405

TAT CAC TTC AAA TCT AAG CAT Tyr His Phe Lys Ser Lys His

66

- (2) INFORMATION FOR SEQ ID NO: 20:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 441 base pairs
- 20 (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
- 25 (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
- 30 (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- 35 (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
- 40 (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:20:FROM 1 TO 441

- 80 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

	ATG	GAG	GTG	GTI	CAG	GAG	GGG									.2
	Met	Glu	Val	Val	Gln	Glu	Gly	•								_
5		655					[,] 660									
	CCT	GCT	CAG	AAG	GAG	CTG	CTG	CCT	ccc	GTG	GAG	CCT	GCI	CAG	ATG	6
	Pro	Ala	Gln	Lys	Glu	Leu	Leu	Pro	Pro	Val	Glu	Pro	Ala	Gln	Met	
					665					670					675	
10																
	GTG	GGT	GCC	CAA	ATT	GTA	CTT	GCT	CAC	ATG	GAG	CTG	CCT	CCT	ccc	11:
	Val	Gly	Ala	Gln	Ile	Val	Leu	Ala	His	Met	Glu	Leu	Pro	Pro	Pro	
					680					685					690	
15	ATG	GAG	ACT	GCT	CAG	ACG	GAG	GTT	GCC	CAA	ATG	GGG	CCT	GCT	CCC	156
	Met	Glu	Thr	Ala	Gln	Thr	Glu	Val	Ala	Gln	Met	Gly	Pro	Ala	Pro	
					695					700					705	
	N.M.C		~~~													
20	ATG	GAA	CCT	GCT	CAG	ATG	GAG	GTT	GCC	CAG	GTA	GAA	TCT	GCT	CCC	201
20	met	GIU	PIO	Ala	Gln	Met	Glu	Val	Ala	Gln	Val	Glu	Ser	Ala	Pro	
					710					715					720	
	ATG	CAG	GTG	GTC	CNC	ממכ	CNC	aam	amm							
	Met	Gln	Val	Val	CAG Gln	Tare	GAG	Dec	GTT	CAG	ATG	GAG	CTG	TCT	CCT	246
25			• • • • • • • • • • • • • • • • • • • •		725	Lys	GIU	PIO	val		Met	GIn	Leu	Ser		
					,,,					730					735	
	ccc	ATG	GAG	GTG	GTC	CAG	AAG	GAG	רכזי	CTT	ראכ	מיד מ	CNC	CTC	mom.	203
					Val											291
					740					745	0 111	116	GIU	Deu	750	
30															730	
	CCT	CCC	ATG	GAG	GTG	GTC	CAG	AAG	GAA	CCT	GTT	AAG	ATA	GAG	CTG	336
					Val											550
					755					760	-				765	
35	TCT	CCT	CCC	ATA	GAG	GTG	GTC	CAG	AAG	GAG	CCT	GTT	CAG	ATG	GAG	381
					Glu											_
					770					775					780	

- 81 -

TTG TCT CCT CCC ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG 426 Leu Ser Pro Pro Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg 785 790 795

5 GAG CCA CCT CCT CCC

441

Glu Pro Pro Pro Pro

800

- (2) INFORMATION FOR SEQ ID NO: 21:
- 10 (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 15 (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
- 20 (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-
- 25 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 30 (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:21:FROM 1 TO 48
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATG GAG GTG GTT CAG GAG GGG Met Glu Val Val Gln Glu Gly

21

655

CCT GCT CAG AAG GAG CTG CTG CCT CCC
Pro Ala Gln Lys Glu Leu Leu Pro Pro
665

48

- 5 (2) INFORMATION FOR SEQ ID NO: 22:
 - (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
- 10 (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
- 15 (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
- 20 (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
- 25 (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
- 30 (K) RELEVANT RESIDUES IN SEQ ID NO:22:FROM 1 TO 48 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATG CAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG CTG TCT CCT 45 Met Gln Val Val Gln Lys Glu Pro Val Gln Met Glu Leu Ser Pro

35 725 730 735

CCC 48

Pro

- 40 (2) INFORMATION FOR SEQ ID NO: 23:
 - (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 48 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA to mRNA
- 5 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
- 10 (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
- 15 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
- 20 (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:23:FROM 1 TO 48
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATG GAG GTG GTC CAG AAG GAG CCT GTT CAG ATA GAG CTG TCT

Met Glu Val Val Gln Lys Glu Pro Val Gln Ile Glu Leu Ser

740

745

750

30 ccr ccc

48

Pro Pro

- (2) INFORMATION FOR SEQ ID NO: 24:
- (i) SEQUENCE CHARACTERISTICS
- 35 (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
- 40 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human
- (H) CELL LINE: HeLa
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
- 5 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts 10 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
- 15 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:24:FROM 1 TO 48 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
- ATG GAG GTG GTC CAG AAG GAA CCT GTT AAG ATA GAG CTG

 39

 Met Glu Val Val Gln Lys Glu Pro Val Lys Ile Glu Leu

 755

 760

 765

TCT CCT CCC

Ser Pro Pro

- (2) INFORMATION FOR SEQ ID NO: 25:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
- 30 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 35 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
- 40 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,

Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail

- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 5 (D) VOLUME: BO
 - (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:25:FROM 1 TO 48
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATA GAG GTG GTC CAG AAG GAG CCT GTT CAG ATG GAG

Ile Glu Val Val Gln Lys Glu Pro Val Gln Met Glu

770 780

15

TTG TCT CCT CCC
Leu Ser Pro Pro
48

- (2) INFORMATION FOR SEQ ID NO: 26:
- 20 (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 48 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: cDNA to mRNA.
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
- 30 (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
- (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-
- 35 Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
- 40 (D) VOLUME: BO
 - (E) ISSUE:
 - (F) PAGES:

- (G) DATE: March 24, 1995
- (K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 48
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
- 5 ATG GGG GTG GTT CAG AAG GAG CCT GCT CAG AGG
 Met Gly Val Val Gln Lys Glu Pro Ala Gln Arg
 785 790 795

GAG CCA CCT CCT CCC

48

10 Glu Pro Pro Pro Pro

800

- (2) INFORMATION FOR SEQ ID NO: 27:
- (i) SEQUENCE CHARACTERISTICS
- 15 (A) LENGTH: 1461 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
- 20 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
- 25 (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: cDNA
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler,
- 30 Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
 - (B) TITLE: REST: A Mammalian Silencer Protein that Restricts Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
- 35 (E) ISSUE:
 - (F) PAGES:
 - (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 1461
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

			C GC													15
	Let	ı Al	a :Al	a Pr		n										
					45											
·5	CIT	TA	T AT	G CT	G GC	A AAT	GTO	GCC	TT	A AC	T GG	G GAZ	GT:	ממם	T GGC	: 60
	Let	.11	e Me	t Le	. Al	a Asr	va]	Ala	a Lei	ı Th	r Gl	/ Glu	. Val	l Ası	n Gly	. 60
					50					55	•				60	
10	AGC	TG	TG	r GAT	TAC	CTG	GTC	GGI	GAZ	GAJ	A AGA	CAG	ATO	GC/	A GAA	105
10	261	Cys	s Cys	s Asp	туз	r Leu	Val	Gly	Glu	Gli	ı Arg	g Gln	Met	Ala	a Glu	
					65					70					75	
	CTG	ATO	CCG	GTI	GGG	GAT	AAC	AAC	نسلسل	י ייירי צ	י מט	י אכיתי	C 3 3	C 2 1	GGA	
	Leu	Met	Pro	Val	Gly	/ Asp	Asn	Asn	Phe	Ser	. Asp	Ser	Glu	GAL	Gly	150
15					80					85					90	
	GAA	GGA	CTI	'GAA	GAG	TCT	GCT	GAT	ATA	AAA	GGT	GAA	CCT	CAI	GGA	195
	GIU	GIY	Leu	Glu	Glu	Ser	Ala	Asp	Ile	Lys	Gly	Glu	Pro	His	Gly	
20					95					100					105	
	CTG	GAA	AAC	ATG	GAA	CTG	AGA	ልርጥ	المالية المالية	C2 2	Omc.					
•	Leu	Glu	Asn	Met	Glu	Leu	Arg	Ser	Leu	GAA	ten	AGC	GTC	GTA	GAA	240
					110		- 3			115	Deu	361	VAI	vaı	120	
25	CCT	CAG	CCT	GTA	TTT	GAG	GCA	TCA	GGT	GCT	CCA	GAT	ATT	TAC	AGT	285
	Pro	Gln	Pro	Val	Phe	Glu	Ala	Ser	Gly	Ala	Pro	Asp	Ile	Tyr	Ser	
					125					130					135	
	TCA	AAT	AAA	CCT	עייניט	GCC	~~ *	C								
30	Ser	Asn	Lys	Ala	Leu	GCC Ala	Pro	GAA	ACA	CCT	GGA	GCG	GAG	GAC	AAA	330
			•		140			GIU	1111	145	GIY	ALA	GIU	Asp		
										-1.5					150	
	GGC	AAG	AGC	TCG	AAG	ACC	AAA	CCC	TTT	CGC	TGT	AAG	CCA	TGC	CAA	375
	Gly	Lys	Ser	Ser	Lys	Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	
35					155					160					165	
	TT N TT	C	000													
	Tur	GAA Gl··	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC 2	ATC	AGA	GTT	420
	- 7 -	JIU	nia		Ser 170	Glu	GIA .	Gln			His	His :	Ile	Arg	Val	
40					1,0					175					180	

	CAC	AGT	GCT	AAG	AAA	TTT	TTT	GTG	GAA	GAG	AGT	GCA	GAG	AAG	CAG	-465
	His	Ser	Ala	Lys	Lys	Phe	Phe	Val	Glu	Glu	Ser	Ala	Glu	Lys	Gln	
					185					190					195	
5	GCA	AAA	GCC	AGG	GAA	TCT	GGC	TCT	TCC	ACT	GCA	GAA	GAG	GGA	GAT	510
	Ala	Lys	Ala	Arg	Glu	Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
					200					205					210	
	TTC	TCC	AAG	GGC	CCC	ATT	CGC	TGT	GAC	CGC	TGC	GGC	TAC	AAT	ACT	555
10	Phe	Ser	Lys	Gly	Pro	Ile	Arg	Cys	Asp	Arg	Cys	Gly	Tyr	Asn	Thr	
					215					220			•		225	
			TAT													600
	Asn	Arg	Tyr	Yab	His	Tyr	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	
15					230					235					240	
			GAT													645
	Ala	Gly	Asp	Asn	Glu	Arg	Val	Tyr	Lys	Cys	Ile	Ile	Cys	Thr	Tyr	
20					245					250					255	
20																
			GTG													690
	Thr	Thr	Val	Ser		Tyr	His	Trp	Arg	Lys	His	Leu	Arg	Asn	His	
					260					265					270	
25		222														
23			AGG													735
	Pne	PIO	Arg	Lys		lyr	Thr	Cys	Gly		Cys	Asn	Tyr	Phe		
					275					280					285	
	CNC	ארא	222	B B C	א א א א	~~ ~	~mm	010	~~~							
30			AAA													780
50	vah	Arg	Lys	MSII	290	ıyı	val	GIN	HIS			Inr	HIS	Thr	•	
					290					295					300	
	GAA	CGC	CCA	ጥልጥ	מממ	ጥርጥ		CTMT	TO T	CCT	ריים כיים	TC N	N.C.	mcm.	C2.C	005
			Pro													825
35	.	AL 9	110	- y -	305	Cys	GIU	Deu	Cys		ıyı	261	Sei	Ser		
-					303					310					315	
	AAG	ACT	CAT	СТЪ	ACT	AGA	ፐፋጋ	באדכ	ССТ	ልርጥ	ר אַ יי	ጥር አ	CCT	GNG	ממ	870
			Hiş													570
	-1-				320	3			~~ 9	325		JU1	G L y	3 + u	330	
10										323					330	

	CC	A TT	T AA	A TG	T GA	CAC	TGC	AGI	TA	r GT	G GCC	TCI	' AA'	ר רא:	ם כמי	r 915
	.Pro	o Ph	e Ly:	з Су	s Ası	Glr	ı Cys	Ser	туз	r Vai	l Ala	Ser	Ası	a Gla	n His	. ,,,
					335				_	340					345	
5	GA.	A GT	A ACC	CG	CAT	CC	AGA	CAG	GTT	CAC	TAA C	. GGG	CCI	LAA 1	a cci	960
•	Gli	ı Va	l Thi	Arg	J His	: Ala	Arg	Gln	Val	His	a Asn	Gly	Pro	Lvs	Pro)
					350					355		•		-	360	
	CTI	CAA	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	. AAC	1005
10	Leu	Asr	Cys	Pro) His	Cys	Asp	Tyr	Lys	Thr	Ala	Asp	Arq	Ser	Asn	
					365					370		-		'-	375	
	TTC	AAA :	AAA	CAI	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1050
	Phe	Lys	Lys	His	Val	Glu	Leu	His	Val	Asn	Pro	Arg	Gln	Phe	Asn	
15					380					385		_			390	
	TGC	CCI	GTA	TGT	GAC	TAT	GCA	GCT	TCC	AAG	AAG	TGT	AAT	CTA	CAG	1095
	Cys	Pro	Val	Cys	Asp	Tyr	Ala	Ala	Ser	Lys	Lys	Cys	Asn	Leu	Gln	
					395					400		_			405	
20																
	TAT	CAC	TTC	AAA	TCT	AAG	CAT	CCT	ACT	TGT	CCT	AAT	AAA	ACA	ATG	1140
	Tyr	His	Phe	Lys	Ser	Lys	His	Pro	Thr	Cys	Pro	Asn	Lys	Thr	Met	
					410					415					420	
•																
25	GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	GCT	1185
	Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lys	Lys	Arg	Glu	Ala	
					425		•			430					435	
20	GAC	TTG	CCT	GAT	AAT	TTA	ACC	AAT	GAA	AAA	ACA	GAA	ATA	GAA	CAA	1230
30	Asp	Leu	Pro	Asp	Asn	Ile	Thr	Asn	Glu	Lys	Thr	Glu	Ile	Glu	Gln	
					440					445					450	
							•									
	ACA	AAA	ATA	AAA	GGG	GAT	GTG	GCT	GGA	AAG	AAA	TAA	GAA	AAG	TCC	1275
	Thr	Lys	Ile	Lys	Gly	Asp	Val	Ala	Gly	Lys	Lys	Asn	Glu	Lys	Ser	
35					455					460					465	
	GTC	AAA	GCA	GAG	AAA	AGA	GAT	GTC	TCA	AAA	GAG	AAA.	AAG	CCT	TCT	1320
	Val	Lys	Ala	Glu	Lys	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
					470					475					480	
Ю																

AAT AAT GTG TCA GTG ATC CAG GTG ACT ACC AGA ACT CGA AAA TCA 1365 Asn Asn Val Ser Val Ile Gln Val Thr Thr Arg Thr Arg Lys Ser 495

5 GTA ACA GAG GTG AAA GAG ATG GAT GTG CAT ACA GGA AGC AAT TCA 1410 Val Thr Glu Val Lys Glu Met Asp Val His Thr Gly Ser Asn Ser 500 505 510

GAA AAA TTC AGT AAA ACT AAG AAA AGC AAA AGG AAG CTG GAA GTT 1455 10 Glu Lys Phe Ser Lys Thr Lys Lys Ser Lys Arg Lys Leu Glu Val 515 520 525

GAC AGC

Asp Ser

15

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 1284 base pairs

(2) INFORMATION FOR SEQ ID NO: 28:

- (B) TYPE: nucleic acid
- 20 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA to mRNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 25 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Human
 - (H) CELL LINE: HeLa
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: CDNA
- 30 (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Chong, Jayhong A., Tapia-Ramírez José, Toledo-Aral, Juan, Zheng, Yingcong, Boutros, Michael C., Altschuler, Yelena M., Frohman, Michael A., Kraner, Susan D., Mandel, Gail
- (B) TITLE: REST: A Mammalian Silencer Protein that Restricts
- 35 Sodium Channel Gene Expression to Neurons
 - (C) JOURNAL: Cell
 - (D) VOLUME: 80
 - (E) ISSUE:
 - (F) PAGES:
- 40 (G) DATE: March 24, 1995
 - (K) RELEVANT RESIDUES IN SEQ ID NO:26:FROM 1 TO 1284
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

					A GGZ											21
	Ser	Se	Gly	/ Gl	y Gly	, Gly	' Lei	1								
		10					15									
_																
5	TTI	ACC	AGO	AG	r GGC	.AAC	ATT	GGZ	ATG	GCC	CTO	3 007	AA.	C GA	ATG	66
	P	he 1	hr s	er s	Ser G	ly A	sn I	le G	ly M	iet 7	lla I	Leu I	ro l	Asn A	Asp M	et
					20					25					30	
	TAT	GAC	TTG	CAT	GAC	CIT	TCC	AAA	GCI	GAA	CTO	GCC	GC	CCI	CAG	111
10	Tyr	Asp	Leu	His	Asp	Leu	Ser	Lys	Ala	Glu	Lev	Ala	Ala	Pro	Gln	
					35					40			•	•	45	
	CTT -	ATT	ATG	CTG	GCA	AAT	GTG	GCC	TTA	ACT	. GGG	GAA	GTA	LAAT	GGC	156
15	Leu	Ile	Met	Leu		Asn	Val	Ala	Leu	Thr	Gly	Glu	Val	Asn	Gly	
15					50					55					60	
	AGC	TGC	TGT	GAT	TAC	CTG	GTC	GGT	GAA	GAA	AGA	CAG	ATG	GCA	GAA	201
	ser	Cys	Cys	Asp	Tyr	Leu	Val	Gly	Glu	Glu	Arg	Gln	Met	Ala	Glu	
20					65					70					75	
20	-															
	C1G	ATG	CCG	GTT	GGG	GAT	AAC	AAC	TTT	TCA	GAT	AGT	GAA	GAA	GGA	246
	beu	met	PIO	vai	Gly	Asp	Asn	Asn	Phe	Ser	Asp	Ser	Glu	Glu	Gly	
					80					85					90	
25	CNN	CCA		C2.3	~~											
	Glu	GUA	Ton	GAA	GAG	TCT	GCT	GAT	ATA	AAA	GGT	GAA	CCT	CAT	GGA	291
	014	Gry	Deu	GIU	Glu	ser	АТА	Asp	Ile		Gly	Glu	Pro	His	Gly	
					95					100					105	
	CTG	GAA	220	እጥር	GNN	CTC	202	3. Cm								
30	Leu	Glu	Asn	Met	GAA	TAN	AGA	AGT	TTG	GAA	CTC	AGC	GTC	GTA	GAA	336
				.,	Glu 110	⊅eu	Arg	ser	Leu		Leu	Ser	Val	Val		
					110					115					120	
	CCT	CAG	ССТ	АТЭ	TTT	GNG	GCN	ጥሮአ	CCT	com	222					
	Pro	Gln	Pro	Val	Phe	Glu	Ala	Sex	GGI	BCI NIA	CCA	GAT	ATT	TAC	AGT	381
35					125	014	Ala	361	GIY		PIO	Asp	iie	Tyr		
										130					135	
	TCA	TAA	AAA	GCT	CTT	GCC	CCT	ממ	ארייא	CCT		000		~ -		
	Ser	Asn	Lys	Ala	Leu	Ala	Pro	GNA	Th~	Dra	GGA	GCG	GAG	GAC	AAA	426
			-1-		140	44	0	91 4			сту	ATG	GIU	Asp		
40										145					150	

	GGC	AAG	AGC	TCG	AAG	ACC	AAA	ccc	TTT	CGC	TGT	AAG	CCA	TGC	CAA	47
	Gly	Lys	Ser	Ser	Lys	Thr	Lys	Pro	Phe	Arg	Cys	Lys	Pro	Cys	Gln	
					155					160					165	
5															GTT	51
	Tyr	Glu	Ala	Glu	Ser	Glu	Glu	Gln	Phe	Val	His	His	Ile	Arg	Val	
					170					175					180	
															CAG	56:
10	His	Ser	Ala	Lys	Lys	Phe	Phe	Val	Glu	Glu	Ser	Ala	Glu	Lys	Gln	
	•				185					190				-	195	
			GCC													606
	Ala	Lys	Ala	Arg	Glu	Ser	Gly	Ser	Ser	Thr	Ala	Glu	Glu	Gly	Asp	
15					200					205					210	
			AAG													651
	Phe	Ser	Lys	Gly		Ile	Arg	Cys	Asp	Arg	Cys	Gly	Tyr	Asn	Thr	
20					215					220					225	
20																
			TAT													696
	ASI	Arg	Tyr	Asp		Tyr	Thr	Ala	His		Lys	His	His	Thr	Arg	
					230					235					240	
25	CCT	ccc	~ m		63.6	001										
49			GAT													741
	AIG	GIY	Asp	ASII		Arg	vai	туг	Lys		Ile	Ile	Cys	Thr		
					245					250					255	
	ACA	ארא	GTG) AGC	GNG	ጥአጥ	CNC	#CC.	».cc		03 m					
30																786
			Val	361	260	TYT	nis	тър	Arg		HIS	Leu	Arg	Asn		
					200					265					270	
	TTT	CCA	AGG	444	СТЪ	ፐልሮ	ארא	T CT	GGN	תתת	TCC	220	ምንም	alouto-to	Tr.C.N	007
			Arg													831
35				2,5	275	• y =	****	Cys	Gry		Cys	ASII	lyi	Pne		
					2,5					280		_			285	
	GAC	AGA	AAA	חממ	דממ	ጥልጥ	CTT	CAG	ርልጥ	מייים	እሮን	ע הער	ר א תי	אכיא	CCX	076
			Lys													876
		J	_, _,		290	- 7 -	- 41		.1.4.3	295	AT 9	1111	112	THE	300	
40										233					300	

	GAA	CGC	CCA	TAT	AAA	TGT	GAA	CTI	TGI	CCI	TAC	TCA	AGT	TCI	CAC	921
	Glu	Arg	Pro	Tyr	Lys	Cys	Glu	Leu	Cys	Pro	Туг	Ser	Ser	Ser	Glr	1
					305					310					315	
_																
5		ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	ACT	CAT	TCA	GGT	GAG	AAG	966
	Lys	Thr	His	Leu	Thr	Arg	His	Met	Arg	Thr	His	Ser	Gly	Glu	Lys	
					320					325					330	
10	CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	1011
10	Pro	Phe	Lys	Cys	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
					335					340				`-	345	
				_												
	GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	1056
15	GIU	val	Thr	Arg		Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lys	Pro	
13					350					355					360	
	Ton	AAI	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	1101
	Deu	ASII	cys	PIO		Cys	Asp	Tyr	Lys	Thr	Ala	Asp	Arg	Ser	Asn	
20					365					370					375	
20	مانلسك	מממ	מממ	Cam	Cm s	G) G										
	Phe	Lve	Tue	UAI	GIA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1146
		Lys	Dys	nis	380	GIU	геп	HIS	Val		Pro	Arg	Gln	Phe		
					360					385					390	
25	TGC	ССТ	GTA	тст	GAC	ידמיד	CCN	CCT	maa							1191
	Cys	Pro	Val	Cvs	Asp	Tvr	Ala	מל מ	100	AAG	AAG	Cys	AAT	CTA	CAG	1191
	•			-,-	395	-1-	A14	AIG	SEL	400	rys	Cys	ASN	Leu		
										400					405	
	TAT	CAC	TTC	AAA	TCT	AAG	САТ	CCT	א ריידי	тст	CCT	ידית מ	אאא	202	B.T.C	1236
30	Tyr	His	Phe	Lys	Ser	Lvs	His	Pro	Thr	Cve	Pro	Asn	THE	MLA The	Man	1236
	-			-	410					415	110	A311	Lys	1111		
															420	
	GAT	GTC	TCA	AAA	GTG	AAA	CTA	AAG	AAA	ACC	AAA	AAA	CGA	GAG	دس	1281
	Asp	Val	Ser	Lys	Val	Lys	Leu	Lys	Lys	Thr	Lvs	Lys	Ara	Glu	Δla	1201
35					425	-		-	•	430	-,-	,	=		435	
										·						
	GAC															1284
	Asp															

- 40 (2) INFORMATION FOR SEQ ID NO: 29:
 - (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 28 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- 5 (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: rat
- (vii) IMMEDIATE SOURCE:
- 10 (A) LIBRARY: Genomic
 - (x) PUBLICATION INFORMATION:
 - (A) AUTHORS: Maue, R.A., Kraner, Goodman, R.H., Mandel, Gail
- (B) TITLE: REST: Neuron-Specific Expression of the Rat Brain Type II Sodium Channel Gene Is Directed by Upstream Regulatory
 15 Elements
 - (C) JOURNAL: Neuron
 - (D) VOLUME: 4
 - (F) PAGES: 223-231
- 20 (G) DATE: February, 1990
 - (K) RELEVANT RESIDUES IN SEQ ID NO:29:FROM 1 TO 28
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATTGGGTTTC AGAACCACGG ACAGCACC

- (2) INFORMATION FOR SEQ ID NO: 30:
- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
- 30 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: Genomic DNA
 - (iii) HYPOTHETICAL: no
 - (iv) ANTI-SENSE: no
- 35 (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Rat
 - (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: Genomic
 - (x) PUBLICATION INFORMATION:
- 40 (A) AUTHORS: Maue, R.A., Kraner, Goodman, R.H., Mandel, Gail
 - (B) TITLE: REST: Neuron-Specific Expression of the Rat Brain Type II Sodium Channel Gene Is Directed by Upstream Regulatory

- 95 -

Elements

- (C) JOURNAL: Neuron
- (D) VOLUME: 4
- 5 (F) PAGES: 223-231
 - (G) DATE: February, 1990
 - (K) RELEVANT RESIDUES IN SEQ ID NO:30:FROM 2353 TO 2400
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
- 10 ATTGGGGGGA CGAACCACGG ACAGCACC

What is claimed is:

- 1 1. A substantially pure nucleic acid comprising a nucleic acid encoding a protein
- 2 having at least about 85% homology to at least the DNA binding domain or the suppressor
- 3 domain of an animal REST protein.
- 1 2. The substantially pure nucleic acid of claim 1, comprising a nucleic acid encoding
- 2 at least the DNA binding domain or the suppressor domain of an animal REST protein.
- 1 3. The substantially pure nucleic acid of claim 2, wherein the REST protein is a
- 2 mammalian REST protein.
- 1 4. The substantially pure nucleic acid of claim 3, wherein the REST protein is a
- 2 human REST protein.
- 1 5. The substantially pure nucleic acid of claim 4, wherein the nucleic acid comprises
- 2 SEQ ID NO:2.
- 1 6. The substantially pure nucleic acid of claim 5, wherein the nucleic acid comprises
- 2 SEQ ID NO:10.
- 1 7. The substantially pure nucleic acid of claim 1, comprising a nucleic acid encoding
- 2 both the DNA binding domain and the suppressor domain of an animal REST protein.
- 1 8. The substantially pure nucleic acid of claim 7, wherein the REST protein is a
- 2 mammalian REST protein.
- 1 9. The substantially pure nucleic acid of claim 8, wherein the REST protein is a
- 2 human REST protein.
- 1 10. The substantially pure nucleic acid of claim 9, wherein the nucleic acid comprises
- 2 SEQ ID NO:2.

11. The substantially pure nucleic acid of claim 10, wherein the nucleic acid comprises .1 2 SEQ ID NO:10. 12. The substantially pure nucleic acid of claim 1, comprising a nucleic encoding a 1 2 protein differing from an animal REST protein by no more than about 20 point mutations. 13. A substantially pure nucleic acid that hybridizes with an animal REST nucleic acid 1 2 under stringent conditions. 14. The substantially pure nucleic acid of claim 13, comprising the nucleic acid of 1 2 SEQ ID NO:1. 1 15. A substantially pure nucleic acid comprising a nucleic acid encoding a protein that 2 binds to a promoter having at least about 90% homology to nucleotides 6-28 of SEQ ID NO:29 3 and acting to suppress the activity of a promoter having said promoter. 1 16. A substantially pure protein having at least about 85% homology with at least the 2 DNA binding domain or the suppressor domain of an animal REST protein. 1 17. The substantially pure protein of claim 16, comprising at least the DNA binding domain or the suppressor domain of an animal REST protein. 1 18. The substantially pure protein of claim 17, comprising the protein of SEQ ID 2 NO:2. 1 19. The substantially pure protein of claim 18, comprising both the DNA binding domain and the suppressor domain of an animal REST protein. 20. The substantially pure protein of claim 19, comprising the protein of SEQ ID 1 2 NO:10. 3

- 1 21. A transformed eukaryotic or prokaryotic cell comprising a nucleic acid encoding a 22 protein having at least about 85% homology to at least one of the DNA binding domain or the 3 suppressor domain of an animal REST protein.
- 1 22. The transformed cell of claim 21 comprising a nucleic acid encoding at least the 2 DNA binding domain or the suppressor domain of an animal REST protein.
- 1 23. The transformed cell of claim 22, wherein the REST protein is a mammalian 2 REST protein.
- 1 24. The transformed cell of claim 23, wherein the REST protein is a human REST protein.
- 1 25. The transformed cell of claim 24, wherein the nucleic acid comprises SEQ ID 2 NO:2.
- 26. A vector capable of reproducing in a eukaryotic or prokaryotic cell comprising a nucleic acid encoding a protein having at least about 85% homology to at least the DNA
- 3 binding domain or the suppressor domain of an animal REST protein.
- 1 27. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 26, 2 comprising a nucleic acid encoding at least the DNA binding domain or the suppressor domain
- 3 of an animal REST protein.
- 28. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 27, wherein the REST protein is a mammalian REST protein.
- 1 29. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 28, 2 wherein the REST protein is a human REST protein.
- 1 30. The vector capable of reproducing in a eukaryotic or prokaryotic cell of claim 29, wherein the nucleic acid comprises SEQ ID NO:2.

- 31. A method of preparing a protein having REST activity, wherein the protein has at least about 85% homology with at least the DNA binding domain or the suppressor domain of an animal REST protein, the method comprising:
- 4 (a) transforming an appropriate eukaryotic or prokaryotic cell with an
 5 expression vector for expressing intracellularly or extracellularly a nucleic acid encoding the
 6 protein;
- 7 (b) growing the transformed cell in culture; and
- 8 (c) isolating the protein from the transformed cell or the culture medium.
- 32. A pharmaceutical composition for treating an animal having de-differentiated neural cells or neural cells exhibiting diminished activity comprising an effective amount of a
- REST-interfering nucleic acid, wherein the REST-interfering nucleic acid comprises an
- 4 antisense molecule directed against REST expression or an expression vector for expressing
- 5 REST DNA binding activity but not REST silencer activity, and a pharmaceutically acceptable 6 carrier.
- 1 33. The pharmaceutical composition of claim 32, wherein the animal has brain cancer.
- 1 34. The pharmaceutical composition of claim 32, wherein said animal has a
- 2 demyelinating myasthenia gravis, muscular dystrophy, botulism, peripheral neuropathies.
- 3 traumatic nerve injury, post stroke degeneration, post-traumatic spinal and neural degeneration,
- 4 poliomyelitis or rabies.
- 1 35. A pharmaceutical composition for an animal having neural cells exhibiting
- 2 excessive neural activity comprising an effective amount of an expression vector comprising a
- 3 nucleic acid encoding a protein that inhibits the expression of neural proteins in non-neural
- 4 tissues, and a pharmaceutically acceptable carrier.
- 1 36. The pharmaceutical composition of claim 35, wherein the animal has epilepsy,
- 2 Lennox-Gastaut syndrome, spasticity, trauma-induced pain, schizophrenia, stroke or a
- 3 neurodegenerative disease.

1 37. The pharmaceutical composition of claim 36, wherein the animal has Alzheimer's, 2 Parkinson's or Huntington's disease. 1 38. The pharmaceutical composition of claim 36, wherein the animal has epilepsy. 39. The pharmaceutical composition of claim 36, wherein the animal has a 1 2 neurodegenerative disease. 40. A method of determining the level of REST expression in a tissue sample 1 2 comprising: 3 contacting the tissue sample with (i) a nucleic acid that binds to REST (a) mRNA under stringent conditions or (ii) an antibody specific for REST; 4 5 washing the tissue sample to remove non-specific hybridizations of the 6 nucleic acid or non-specific antibody binding; and 7 determining the level of hybridized nucleic acid or bound antibody. (c) 41. An antibody that reacts specifically with the substantially pure protein of claim 16. 1 1 42. A pair of PCR primers capable of directing the amplification of the substantially pure nucleic acid of claim 1.

Fig. 1 · (Part 1 of 6)

						IGTIC								-275	
						GCCGC								-225	
GCGG	CGGC	TG	CGGC	AGCCC	SA GA	ACGGC	CAGGO	G CGA	AGGC	CCGG	AGG	CCTG	AGC	-175	
ACCC	TCT	CA C	SCCC	CACTO	C TO	GGCC	TTCT	TGO	STCC	ACGA	CGG		AGC	-125	
ACCC	AACT	TT A	ACCAC	CCT	CC CC	CAC	CTCT	ccc	CGA	AACT	CCA	CAA	CAA	-75	
AGAA	LAAGI	AG :	rcgg?	GAAC	G AC	GCGG	GACT	CAC	GGT	CGCC	CGC	CCT	CT	-25	
CACC	GAGG	AA C	GCCG	SAATA	AC AC	GTT								-1	
ATG	GCC	ACC	CAG	GTA	ATG	GGG	CAG	TCT	TCT	GGA	GGA	GGA	GGG	CTG	45
Met	Ala	Thr	Gln	Val	Met	Gly	Gln	Ser	Ser	Glv	Glv	Glv	Glv	Leu	
1				5		•			10	-	3	4	2	15	
TTT	ACC	AGC	AGT	GGC	AAC	ATT	GGA	ATG	GCC	CTG	CCT	AAC	GAC	ATG	90
Phe	Thr	Ser	Ser	Gly	Asn	Ile	Gly	Met	Ala	Leu	Pro	Asn	Asp	Met	
				20			•		25			_		30	
TAT	GAC	TTG	CAT	GAC	CTT	TCC	AAA	GCT	GAA	CTG	GCC	GCA	CCT	CAG	135
Tyr	Asp	Leu	His	Asp	Leu	Ser	Lvs	Ala	Glu	Leu	Ala	Ala	Pro	Gln	
				35			•		40					45	
CTT	ATT	ATG	CTG	GCA	AAT	GTG	GCC	TTA	ACT	GGG	GAA	GTA	AAT	GGC	180
Leu	Ile	Met	Leu	Ala	Asn	Val	Ala	Leu	Thr	Glv	Glu	Val	Asn	Glv	
				50					55	1				60	
AGC	TGC	TGT	GAT	TAC	CTG	GTC	GGT	GAA		AGA	CAG	ATG	GCA		225
Ser	Cys	Cys	Asp	Tyr	Leu	Val	Glv	Glu	Glu	Arg	Gln	Met	Ala	Glu	
	•	•	•	65					70	5				75	
CTG	ATG	CCG	GTT	GGG	GAT	AAC	AAC	TTT	TCA	GAT	AGT	GAA	GAA		270
Leu	Met	Pro	Val	Gly	Asp	Asn	Asn	Phe	Ser	Asp	Ser	Glu	Glu	Glv	
				80	•				85					90	
GAA	GGA	CTT	GAA	GAG	TCT	GCT	GAT	ATA		GGT	GAA	CCT	CAT		315
Glu	Gly	Leu	Glu	Glu	Ser	Ala	Asp	Ile	Lvs	Glv	Glu	Pro	His	Glv	
	-			95			•		100					105	
CTG	GAA	AAC	ATG	GAA	CTG	AGA	AGT	TTG	GAA	CTC	AGC	GTC	GTA		360
Leu	Glu	Asn	Met	Glu	Leu	Arg	Ser	Leu	Glu	Leu	Ser	Val	Val	Glu	
				110		_			115					120	
CCT	CAG	CCT	GTA	TTT	GAG	GCA	TCA	GGT	GCT	CCA	GAT	ATT	TAC		405
						Ala									
				125					130				- 2 -	135	
TCA	AAT	AAA	GCT	CTT	GCC	CCT	GAA	ACA	CCT	GGA	GCG	GAG	GAC		450
Ser	Asn	Lys	Ala	Leu	Ala	Pro	Glu	Thr	Pro	Glv	Ala	Glu	Asp	Lvs	
		-		140					145					150	
GGC	AAG	AGC	TCG	AAG	ACC	AAA	CCC	TTT	CGC	TGT	AAG	CCA	TGC		495
						Lys									
-	•			155	•	•			160					165	
TAT	GAA	GCA	GAA	TCT	GAA	GAA	CAG	TTT	GTG	CAT	CAC	ATC	AGA		540
						Glu									
				170					175					180	
CAC	AGT	GCT	AAG	AAA	TTT	TTT	GTG	GAA	GAG	AGT	GCA	GAG	AAG	CAG	585
						Phe									
			_	185					190				-	195	
GCA	AAA	GCC	AGG	GAA	TCT	GGC	TCT	TCC	ACT	GCA	GAA	GAG	GGA		630
						Gly									
	-			200		_			205				-	210	

Fig. 1 Part 2 of 6

TTC	TCC	AAG	GGC	CCC	ATT	CGC	тст	GAC	CGC	TGC	GGC	ጥልር	እ አጥ	א כידי	675
Phe	Ser	Lys	Gly	Pro	Ile	Ara	Cvs	ASD	Ara	Cys	Glv	Tvr	yen	Thr	075
		_		215		5			220					.225	
TAAT	CGA	TAT	GAT	CAC	TAT	ACA	GCA	CAC	CTG	AAA	CAC	CAC	ACC	AGA	7.20
Asn	Arq	Tyr	Asp	His	TYT	Thr	Ala	His	Leu	Lys	His	His	Thr	Arg	7.20
				.230					235					240	
GCT	GGG	GAT	.AAT	GAG	CGA	GTC	TAC	AAG	TGT	ATC	ATT	TGC	ACA	TAC	765
Ala	Gly	Asp	Asn	Glu	Arq	Val	Tvr	Lvs	Cvs	Ile	Tle	Cvs	Thr	Tyr	, 45
	-	-		245			- 1 -	-1-	250		110			255	
ACA	ACA	GTG	AGC	GAG	TAT	CAC	TGG	AGG	AAA	CAT	עיוייד	AGA	AAC	CAT	810
Thr	Thr	Val	Ser	Glu	Tyr	His	Trp	Arg	Lvs	His	Leu	Ara	Asn	His	010
				260					265					270	•
TTT	CCA	AGG	AAA	GTA	TAC	ACA	TGT	GGA	AAA	TGC	AAC	TAT	TTT	TCA	855
Phe	Pro	Arg	Lys	Val	Tyr	Thr	Cys	Gly	Lys	Cys	Asn	Tvr	Phe	Ser	
				275					280	_				285	
GAC	AGA	AAA	AAC	AAT	TAT	GTT	CAG	CAT	GTT	AGA	ACT	CAT	ACA	GGA	900
<u>Asp</u>	Arg	Lys	Asn	Asn	Tyr	Val	Gln	His	Val	Arq	Thr	His	Thr	Gly	
				290					295					300	
GAA	CGC	CCA	TAT	AAA	TGT	GAA	CTT	TGT	CCT	TAC	TCA	AGT	TCT	CAG	945
Glu	Arg	Pro	Tyr	Lys	Cys	Glu	Leu	Çys	Pro	Tyr	Ser	Ser	Ser	Gln	
				305					310					315	
AAG	ACT	CAT	CTA	ACT	AGA	CAT	ATG	CGT	ACT	CAT	TCA	GGT	GAG	AAG	990
Lys	Thr	<u> His</u>	Leu	Thr	Arq	<u>His</u>	Met	Arg	Thr	His	Ser	Gly	Glu	Lys	
				320					325					330	
CCA	TTT	AAA	TGT	GAT	CAG	TGC	AGT	TAT	GTG	GCC	TCT	AAT	CAA	CAT	1035
Pro	Phe	Lys	<u>Cys</u>	Asp	Gln	Cys	Ser	Tyr	Val	Ala	Ser	Asn	Gln	His	
				335					340					345	
GAA	GTA	ACC	CGC	CAT	GCA	AGA	CAG	GTT	CAC	AAT	GGG	CCT	AAA	CCT	1080
Glu	Val	Thr	Arg	His	Ala	Arg	Gln	Val	His	Asn	Gly	Pro	Lys	Pro	
				350					355					360	
CIT	AAT	TGC	CCA	CAC	TGT	GAT	TAC	AAA	ACA	GCA	GAT	AGA	AGC	AAC	1125
Leu	Asn	Cys	Pro	<u>His</u>	Cys	Asp	Tyr	Lys		Ala	Asp	Arg	Ser		
mm o				365					370					375	
TIC.	AAA	AAA	CAT	GTA	GAG	CTA	CAT	GTG	AAC	CCA	CGG	CAG	TTC	AAT	1170
Phe	LYS	TAR	Hls	Val	Glu	Leu	His	Val		Pro	Arg	Gln	Phe		
TCC	CCT	CER	mam	380	m >				385	9				390	
Cvc	D~0	GIA	TGT	GAC	TAT	GCA	GCT	TCC	AAG	AAG	TGT	AAT	CTA	CAG	1215
CYS	PIO	val	Cys	395	TYT	Ala	Ala	Ser	<u>Lys</u>	Lys	Cys	Asn	Leu		-
יד איד	CNC	ጥጥ	222		3 B C	C % TT			400					405	
Tire	LAC	Pho	AAA	101	AAG	CAT	CCT	ACT	TGT	CCT	AAT	AAA	ACA	ATG	1260
TAT	UT2	FIIE	TAR	410	LVS	HIS	PIO	Thr		Pro	Asn	Lys	Tnr		
CAT	GTC	TCA	אאה		מממ	CTA	330	***	415			CC 1	C . C	420	7205
Den	Val	Ser	Live	77-1	Luc	LAU	MAG	AAA	ML-	AAA	AAA	CGA N==	GAG	71-	1305
A3D	Val	261	пåэ	425	пåг	חבת	гÀг	ьys		Lys	гÀг	Arg	GIU		
GAC	בית	رسد	ር እጥ		אינים ע	אככ	יתיאא	~ ~ ~	430	3.03	C 2 2	2022	C 3 3	435	1250
Acn	Len	Dra	DCD GWI	yez	TIV	カレし	AA1	CI	AAA	ACA	GAA	AIA	GAA	CAA	1350
чэр	חבת	FIO	wab	440	TTE	TIII.	ASI	GIU		Thr	GIU	TTE	GIU		
									445					450	

Fig. 1 Part 3 of 6

			.= .= =												
ACA	AAA	ATA.	AAA	GGG	GAT	GTG	GCI	GGZ	AAG	AAA	AAT	GAA	AAG	TCC	1395
TIT	rys	TTE	Lys	Gly	Asp	Val	Ala	Gly	/ Lys	Lys	Asn	Glu	Lys	Ser	
				·* 33					450	1				4	
77-7	AAA	GCA	GAG	AAA	AGA	GAT	GTC	TC	AAA A	GAG	AAA	AAG	CCT	TCI	1440
val	цуs	Ala	GIU	Thas	Arg	Asp	Val	Ser	Lys	Glu	Lys	Lys	Pro	Ser	
				~ 1 / U					475					400	
yen	yen vvi	Val	COM	616	ATC	CAG	GTG	ACI	, YCC	AGA	ACT	CGA	AAA	TCA	1485
ASII	W211	VAI	261	485	TTE	GIN	Val	Thr	Thr	Arg	Thr	Arg	Lys	Ser	
				400					797						
Val	Thr	Glu	Val	TAKE	CAG	ATG	GAT	GTG	CAT	ACA	GGA	AGC	AAT	TCA	1530
			Val	500	GIU	Met	Asp	val	His	Thr	Gly	Ser	Asn	Ser	•
GAA	AAA	TTC	AGT	777	A CT	220	7.7.7	100	505	AGG				510	
Glu	Lvs	Phe	Ser	Tive	Thr	Lve	TAG	AGC	AAA	AGG Arg	AAG	CTG	GAA	GTT	1575
				515		Lys	nys	Sei	TAS	Arg	Lys	Leu	Glu		
GAC	AGC	CAT	TCT	TTA	CAT	CCT	راس	CTC	520	~ n m	~~~			525	1620
Asp	Ser	His	Ser	Leu	His	Glv	Pro	Val	yez	Asp	GAG	GAA	TCT	TCA	1620
				220					E 3 E						
ACA	AAA	AAG	AAA	AAG	AAG	GTA	GAA	AGC	777	TCC	מממ	አአጥ	220	540	1665
Thr	Lys	Lys	Lys	Lys	Lys	Val	Glu	Ser	Lvs	Ser	Tare	WWI	WWI	AGI	1002
				242					EED						
CAG	GAA	GTG	CCA	AAG	GGT	GAC	AGC	AAA	CTC	GAG	GAG	דממ	מממ		1710
Gln	Glu	Val	Pro	درب	Gly	Asp	Ser	Lys	Val	Glu	Glu	Asn	Lvs	Tive	1/10
				200					555						
CAA	AAT	ACT	TGC	ATG	AAA	AAA	AGT	ACA	AAG	AAG	AAA	ACT	CTG		1755
GIII	ASI	Inr	Cys	MEC	Lys	Lys	Ser	Thr	Lys	Lys	Lys	Thr	Leu	Lys	
				2,2					E 10 /						
Aen	Tare	I CA	AGI	AAG	AAA	AGC	AGT	AAG	CCT	CCT	CAG	AAG	GAA	CCT	1800
11311	Lys	261	Set	БуS 590	Lys	Ser	Ser	Lys	Pro	Pro	Gln	Lys	Glu	Pro	
GTT	GAG	AAG	CCA	TOT	CCT	CNC	3.000		595	20				600	
Val	Glu	Lvs	GUA	Ser	אום	CAG	ATG	GAC	CCT	CCT	CAG	ATG	GGG	CCT	1845
		-,-	0+3	605	AT a	GIII	met	Asp	Pro	Pro	Gln	Met	Gly		
GCT	CCC	ACA	GAG	GCG	ىسك	CAG	AAG	CCC	610	~~~	~ ·			615	1890
Ala	Pro	Thr	Glu	Ala	Val	Gln	Lve	Clv	D~c	Val	CAG	GIG	GAG	CTG	1890
				020					E7E					C 2 2	
CCA	CCT	CCC	ATG	GAG	CAT	GCT	CAG	ATG	GNG	GGT	GCC	CNC	አጥአ	630	3035
Pro	Pro	Pro	Met	Glu	His	Ala	Gln	Met	Glu	Gly	Al=	CAG	TIA	7~~	1935
				ככם					E / D						
CCT	GCT	CCT	GAC	GAG	CCT	GTT	CAG	ATG	CAC	GTG	GTT	CAG	GAG	~~~	1980
Pro	Ala	Pro	Asp	Glu	Pro	Val	Gln	Met	Glu	Val	Val	Gln	Glu	Glv	1700
				0 D U					<u> </u>						
CCT	GCT	CAG	AAG	GAG	CTG	CTG	CCT	CCC	GTG	GAG	CCT	GCT	CAG	200	2025
PIO	ATA	GIn	Lvs	GIU	Leu	Leu	Pro	Pro	Val	Glu	Pro	Ala	Gln	Met	-
				000					670					675	
Val	Cl	71-	CAA	ATT	GTA	CTT	GCT	CAC	ATG	GAG	CTG	CCT	CCT	CCC	2070
AGT	отЪ	WIG	GIII	TTE	val	теп	Ala	His	Met	Glu	Leu	Pro	Pro	Pro	
				680					685					690	

Fig. 1 Part 4 of 6

'N TOC	CRO														
Mat	CAU	ACT	GCI	CAG	ACC	GAC	GT	GCC	CAZ	ATC	GGG	CCI	GCT	ccc	2115
Met	GIU	Inr	ALa	((3.11)	1.111	์ Glบ	Va]	Ala	a Glr	1 Met	: Gly	Pro	Ala	Pro)
				ככס					700	١				~ ~ ~	
Mor	Cl	. CCI	GCI	CAG	ATG	GAG	GTI	GCC	CAC	GTA	GAA	TCI	GCI	, ccc	2160
MEC	GIU	PIO	ATS	GIII	Met	Glu	Val	. Ala	Gli	ı Val	. Glu	Ser	Ala	Pro	
				/. T U					715						
Met	Cla	Val	77-7	CAG	AAG	GAG	CCI	GTI	CAC	ATC	GAG	CTG	TCT	CCI	2205
1100	OLII	Vai	val	725	LYS	GIU	Pro) Val	GLT	<u>Met</u>	<u>Glu</u>	Leu	Ser	Pro	
CCC	ATG	GAG	СТС	745 CTC	C3 C	880	~~		730	_				735	
Pro	Met	Glu	Val	Val	CAG	AAG	GAG	CCI	GTI	CAG	ATA	GAG	CTG	TCT	2250
		<u> </u>		Val 740	GIII	LVS	GIU	PTC	val	Gln	Ile	<u>Glu</u>	Leu		
CCT	CCC	ATG	GAG	GTG	GTC	CAG	222	CAA	745	-				750	2295
Pro	Pro	Met	Glu	Val	Val	Gla	Tare	Cl	D	GIT.	AAG	ATA	GAG	CTG	2295
				/ 25					750						
TCT	CCT	CCC	ATA	GAG	GTG	GTC	CAG	ממ	GAG	رحم	ىسىت	CBC	8 mc	765	2340
<u>Ser</u>	Pro	Pro	Ile	Glu	Val	Val	Gln	Lvs	Clu	Dro	77-1	CAG	Mor	GAG	2340
				//U					775					500	
TTG	TCT	CCI	CCC	ATG	GGG	GTG	GTT	CAG	AAC	CRC	رك	CCT	CAG	9.00	2385
Leu	Ser	Pro	Pro	Met	Gly	Val	Val	Gln	Lvs	Glu	Pro	Ala	Gla	Ara	2303
				700					700					705	
GAG	CCA	CCT	CCT	CCC	AGA	GAG	CCT	CCC	CTT	CAC	ATG	GAG	CCA	8 mm	2430
GIU	Pro	Pro	Pro	<u> </u>	Arg	Glu	Pro	Pro	Leu	His	Met	Glu	Pro	Ile	2130
				000										0 . 0	
100	AAA	AAG	CCT	CCT	CIC	CGA	AAA	GAT	AAA	AAG	GAA	AAG	TCT	3 3 0	2475
Ser	гÀг	rys	Pro	PLO	Leu	Arg	Lys	Asp	Lys	Lys	Glu	Lys	Ser	Asn	
				0.72					820					005	
Met	Cln	So~	Clu	AGG	GCA	CGG	AAG	GAG	CAA	GTC	CTT	ATT	GAA	GTT	2520
MCL	GIII	SeI	GIU	Arg 830	Ala	Arg	Lys	Glu	Gln	Val	Leu	Ile	Glu	Val	
GGC	ተ ተል	GTG	CCT	030	222	~~ m	9.55		835					840	
Glv	Leu	Val	Dro	Val	AAA	GAT	AGC	TGG	CTT	CTA	AAG	GAA	AGT	GTA	2565
				Val 845	Lys	ASD	ser	Trp	ren	Leu	Lys	Glu	Ser		
AGC	ACA	GAG	GAT		TCA	CCB	CCB	TCB	850	CC2				855	2610
Ser	Thr	Glu	Asp	Leu	Ser	Pro	Pro	Ser	DZ2	D	CIG	CCA	AAG	GAA	2610
				טסס					855					0 7 0	
AAT	TTA	AGA	GAA	GAG	GCA	TCA	GGA	GAC	ממי	מממ	עידיידי	CTC	A A C	870	2655
Asn	Leu	Arg	Glu	Glu	Ala	Ser	Glv	ASD	Gln	Live	Len	Lau	AAC	Th~	2655
				0/2					997					~ ~ =	
GGT	GAA	GGA	AAT	AAA	GAA	GCC	CCT	CTT	CAG	AAA	GTA	GGA	GCA	CBB	2700
Gly	Glu	Gly	Asn	Lys	Glu	Ala	Pro	Leu	Gln	Lvs	Val	Glv	Ala	Glu	2,00
				ロソリ					295					900	
GAG	GCA	GAT	GAG	AGC	CTA	CCT	GGT	CTT	GCT	GCT	AAT	ATC	AAC	CAA	2745
GIU	Ala	Asp	Glu	ser	Leu	Pro	Gly	Leu	Ala	Ala	Asn	Ile	Asn	Glu	
				フレコ					910					015	
202	オレビ	CAT	ATT	TCA	TCC	TCT	GGA	CAA	AAC	TTG	AAT	ACG	CCA	GAG	2790
Jei	111T	uT2	11,5	5 e1	ser	Ser	Gly	Gln	Asn	Leu	Asn	Thr	Pro	Glu	
			i	920					925					930	

Fig. 1 Part 5 of 6

GGT	GAA	ACT	TTA	AAT	GGT	AAA	CAT	CAG	ACT	GAC	AGT	ATA	GTT	TGT	2835
Gly	Glu	Thr	Leu	Asn	Gly	Lys	His	Gln	Thr	Asp	Ser	Tle	Val	CVE	2000
				935					940					OAE	
GAA .	ATG	AAA	ATG	GAC	ACT	GAT	CAG	AAC	ACA	ACA	GAG	יי א א	CTC	ינים ע	.2880
Glu 1	Met	Lys	Met	Asp	Thr	Asp	Gln	Asn	Thr	y.~~	GL	y 22	Tan	MC1	.2000
		4		950				A311	955	Arg	GIU	ASII	ren		
GGT	ATA	AAT	TCA		ىلىشت	CAA	C 3 3	CCN	722	ma>				960	
Glv	Tle	Acn	Sar	Th~	77-7	Class	CAA	CCA	GII	TCA	CCA	ATG	CTT	CCC	2925
Gly		7.511	261	965	VAI	GIU	GIU	Pro	vai	Ser	Pro	Met	Leu		
	ጥሮአ	CCN	CITIA						970					975	
D=0	COM	33-	GIA	GAA	GAA	CGT	GAA	GCA	GTG	TCC	AAA	ACT	GCA	CTG	·2970
Pro	ser	Ala	vai	GIU	GIA	Arg	Glu	Ala	Val	Ser	Lys	Thr	Ala	Leu	
				980					925		:			000	
GCA	TCA	CCL	CCI	GCT	ACA	ATG	GCA	GCA	AAT	GAG	TCT	CAG	GAA	ATT	3015
Ala	Ser	Pro	Pro	Ala	Thr	Met	Ala	Ala	Asn	Glu	Ser	Gln	Glu	Ile	
				775					1 N N C	}				100	5
GAT (GAA	GAT	GAA	GGC	ATC	CAC	AGC	CAT	GAA	GGA	AGT	GAC	CTA	አርጥ	3060
Asp (Glu	Asp	Glu	Gly	Ile	His	Ser	His	Glu	Glv	Ser	ASD	T.en	Ser	5000
				TOT	,				1015	:				102/	1
GAC I	AAC	ATG	TCA	GAG	GGT	AGT	GAT	GAT	ىلىكى	GCA	كالس	רא יי	ccc	COT	3105
Asp :	Asn	Met	Ser	Glu	Glv	Ser	ASD	ASD	Ser	Gly	Tau	Unic	Class	81-	3103
_				1025	5				1030	GTA	пеа	nis	Gry		-
CGG (CCA	GTT	CCA			T)	AGC	A C A	RRR	, 2200	003	220	~	1035	3150
Arg 1	Pro	Val	Pro	Gln	Glu	Ser	Ser	Non	AAA Taa	AAI	BLA	AAG	GAA	GCC	3150
				1040)	Der	261	Arg	Lys	ASI	ALA	гÃг	GIU		
ተጥር (מרא	CTC	222			A B C	CCB	~~ m	1045					1050)
Leu	Ala ef&	Val	Tare	220	אן ה	T	Class	GAT	TTT	GTT	TGT	ATC	TTC	TGT	3195
Leu i		V CL	цуs	1055	Ala	гåг	GIA	Asp	Pne	Val	Cys	Ile	Phe		
CAT (CCT	ىلىبىلى	السلام	707	, , , , , , , , , , , , , , , , , , , ,	CCB		~~	1060					1065	5
GAT (2 ~~	502	Pho	y~~	T	Class	AAA	GAT	TAC	AGC	AAA	CAC	CTC	AAT	3240
Asp A	719	SEI	FILE	1070	, nys	GIA	гÀг	Asp	Tyr	Ser	Lys	His	Leu		
CGC (יד גר	كالملك	لمنت			m» c			1075			•		1080)
CGC (uri Lic	TOU	77-1	WWI	77-3	TAC	TAT	CTT	GAA	GAA	GCA	GCT	CAA	GGG	3285
Arg I	uTS	Deu	val	ASI	.vai	lyr	Tyr	Leu	Glu	Glu	Ala	Ala	Gln	Gly	
C3 C (~ ~ ~	m = = =		1085) 				1090					1095	5
CAG (JAG	TAAT	G AA	AACT"	"I'GAA	CAA	GGTI	TCA	GTTC	TTAC	TT				3326
Gln (
	1097														
TGTA	AGGT	AT A	TTAC	CATTI	T AT	ATTC	ATTI	ATO	ATAG	CAG	ACAA	CCTI	TT		3376
AAGA:	TTGC	TT T	L'AAT'I	CAGTA	T CI	GATO	TTGA	TTI	TTAA	GTG	GCAT	TCTI	TT		3426
CCTT	AGGA	CTI	TTT	ATGT	AT AC	CTGI	TGAI	TGI	TGTG	TAA	ATTT	TAGI	'AA		3476
ATCT	AAGA	GA G	STGTA	CTAP	ry CC	AGCA	GGTA	TCI	GTTA	GCT	TATG	TGTI	'TA		3526
ATTG	TAAA	TA G	AAGC	CTAP	G AI	GGTA	TAAC	AGC	TTTA	TAT	TGCT	TTGI	CC		3576
AGCT	ACAA	CA I	GTCA	TTTT	T TI	CTCC	ATGI	CTI	ATCT	TCC	TGTT	TCAC	TT		3626
TAGT:	TAT	TC I	TCGI	TTTT	TA T	TGAG	ATCI	' ATA	AAAA	ATT	GGCT	TACT	TA		3676
ATAG(CAAA	A TT.	CTTG	AAGA	\mathtt{IT} A_{J}	TGCC	TGCI	TTA	TATA	AAG	TTAG	CACT	TT		3726
AAGA:	$\Gamma T T T$	TT T	TTTP	IGAGA	T GA	GAAG	ACAT	' TTA	AATT	GAA	GAAA	AATT	CC		3776
CCCA	GCAA	TA G	AÇAC	TCTA	T CA	GTCC	AAGT	' ATT	TACT	TCC	TGAG	$\frac{1}{2}$	GA		3826
TCAA	TATT	TT T	TATT	TGTG	TAT	GTTA	ATCG	TCA	TAAA	חממ	ACTO	ጉተተ ተ	TC		3876
GTGT	GTTT	TT T	'ATTI	TGGT	G CT	TTAA	TGGC	בדד י	ACAT	شىك م	CCVC	ው መመመ ተጽሞ ተሞ	TO		3926
TTTTT	CTT	TT G	GTTT	CTGT	T TA	TGTT	TTTT	ַ װָבָרָ הַבָּרָ	ית בייטרי	CCA	CTTA	ያያው ተተተተ	ىلىك + +		
TCCT	AGAĀ	AT A	GCAT	TTGT	G TT	CAAC	מתטע.	200		ע תיינ גייי ע	CYWY	መጀመን የጉርፈት ነ	ፈነ ጉጉ		3976
		••								UTW	CWIN	TWIN	TH		4026

WO 96/29433 PCT/US96/03940

Fig. 1 Part 6 of 6

TGCATGTTTA	TTTTGTTTGG	CGTCTTTGGA	GGGATGCTTT	TAGACTTGTT	4076
TGCAAAAGGG	CAGTTTTCTT	TTTCTTTGCT	GCAGTTGTCT	ATTTTGCAGA	4126
ATAATAGTGT	GTGCAAGTTT	GTGAGCAAAT	GAAATATGCA	GGTTCAATCT	4176
ATTGATTTTG	ATTTTTACAT	CTTATATCTA	TGCCAGAATC	TGTATTTCAT	4226
ATAACTTATT	TATTTCGAAT	GGATGTAGTA	AATTCACAGC	TATCAGTTTT	4276
GATTTTGCAA	TAAATAAACC	ACTAGGTTGC	ATGTCGAACA	AATTTTTATC	4326
TCAAATACCA	ACCATCAGTT	TTTTTTTCA	TGTGTTTTGG	TACAGCTAAT	4376
TCCTAATTGT	AGAGTGTTAA	ATGTTTGAGG	AGAACCTTTT	CTCATAGATG	4426
GTTGGTGTTC	ATATGGCNAC	TTTACAATAA	AGAGAACTGT	AAGTGATATT	4476
TGGAAACTAC	AAACCTGGAA	TTAGGAGATA	TAATTATTCC	TTCAAGTTTT	4526
ATAGATATCA	CTTGGGAGAT	TCCAAAGCCA	TAGCTATTAC	GCNGCAAACC	4576
TAGGATAAGA	AAGGTAGTAT	GAGTGCTGGT	AGACCAGCTG	CAACATTTCC	4626
TATATCAGAT	GAAAAAGGCT	GGTGAAACAA	GTACAGTCCA	GATTTTTTAA	4676
AATCATACTT	TCTCAGGGAT	CTCCACAAAC	TGGTGGGTGT	CCTGGCTGTC	4726
TGTGTGATAG	CCTCTTTCTA	TAGGTGAGGC	CTCAAATGAA	TTGCAGCTAT	4776
CCTGGTGTTC	CTATGAGGGC	ACTTGTATGA	AAAAGGCAGT	ACTCCAAAAC	4826
ATTTTTGATG	GTTCTTTGGC	CAGTTGCCAA	AGAGTGTGAA	AGAATCCAAT	4876
AGAGGATTTT	TCTTACTGAT	AGCAGTCATT	CATTGCAGTA	AAATAAAATA	4926
TGAATTCCCA	TTAGGGAATC	TTGAATTCTG	ACCTCCCATA	CTCCGTTTTG	4976
AAATAACCAC	TTATATTTCA	TTTTTTAAAA	ATCTGATGAT	CTCTTTGAGG	5026
CAGGTTTCAG	ATTTGGCAGT	ACAACATGAA	AGATTAGGAA	AAGCATTAAT	5076
AACGTGTGGG	TGGAAAGCTT	GTTAAAAATC	TGAGAGTGAA	GTTTGAGTTA	5126
AAAGTTGTTT	GACATGGCAT	TGACTGGGAG	GCCAAAGATT	TAAAGAAGCG	5176
GAAGATTCTT	CTCTTAAGAC	ATGAGGAGTA	AGTTGTGTGA	TAATGGTATG	5226
TGTTTTGTGT	GCATGAATGG	ACATTGTAAA	TGTTGAATTC	TAGGCTCCGA	5276
CAATCATTGT	CAACAGAAGA	TAAAGCTGCA	AATATTTATG	TTTTAAAA	5324

INTERNATIONAL SEARCH REPORT

International application No. PCT/US96/03940 /

	ASSIFICATION OF SUBJECT MATTER		
IPC(6)	:Please See Extra Sheet.		
According	:435/6, 91.2, 7.1, 7.21, 7.23; 536/ 23.1, 24.3; 530 to International Patent Classification (IPC) or to bot	J/350, 388.2 h national obseitantion and IBC	
	LDS SEARCHED	iii iiatioliai classification and IPC	
	documentation searched (classification system follow	and has almost Counting and had a	
		•	
0.3	435/6, 91.2, 7.1, 7.21, 7.23; 536/ 23.1, 24.3; 530,	/350, 388.2	
Documenta	tion searched other than minimum documentation to the	he extent that such documents	
		ne extent that such documents are included	in the fields searched
Electronic o	data base consulted during the international search (r	name of data base and where provide his	
Please S	ee Extra Sheet.	or dam base and, where practicable	, search terms used)
. DOC	CUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where a	ppropriate, of the relevant passages	Relevant to claim No.
Κ	Science, Volume 267, issued	03 March 1995 C I	1-41
-	Schoenherr et al, "The Neuron-F	Restrictive Silencer Factor	1-41
′	(NRSF): A Coordinate Repressor of	42	
	Genes". Figures 1-6, see entire do	ocument.	72
	-		
′	Henry A. Erlich, "PCR Technology"	', published 1992, by W.H.	42
	Freeman and Co. (N.Y.), pages 7-	16, especially pages 8-10.	
	_	, , , , , , , , , , , , , , , , , , , ,	
			•
	,		
Furth	er documents are listed in the continuation of Box C	See patent family annex.	
=	exial categories of cited documents:		
A* doc	cunent defining the general state of the art which is not considered	date and not in conflict with the applica	tion but cated to understand the
to l	be of particular relevance	principle or theory underlying the invent. 'X' document of puricular relevance: the	
	lier document published on or after the international filing date	considered novel or cannot be consider	ed to involve an inventive step
cite	turnent which may throw doubts on priority chan(s) or which as at to establish the publication date of whother clustion or other distributions (controlled).	when the document is taken alone	
•	cial reason (as specified) cament referring to an oral disclusure, ase, exhibition or other	considered to involve an inventive	step when the document is
ne me	and therefore we all of all disciplines, use, exhibition of other	combined with one or more other such being obvious to a person skilled in th	documents, such combination
	nument published prior to the international filing date but later than priority date claimed	*&* document member of the same patent	funily
	actual completion of the international search	Date of mailing of the international sea	rch report
27 MAY 1	1996	14 JUN 1996	
ame and m	nailing address of the ISA/US	Authorized officer \ \ \(\lambda \) \(\lambda \)	
	ner of Palents and Tradenurks	Addition Liver Liver	Tula /
	, D.C. 20231	DIANNE REES	, , , , ,
acsimile No	o. (703) 305-3230	Telephone No. (703) 308-0196	

(703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No. PCT/US96/03940

IPC (6):

C12Q 1/68; C12P 19/34, 21/08; G01N 33/53, 33/567, 33/574; C07H 21/04; C07K 1/00

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, BIOTECHABS, BIOTECHDS, CAPLUS, CABA, CANCERLIT, DISSABS, DGENE, DRUGU, EMBASE, GENBANK, PROMT, TOXLINE, TOXLIT, USPATFULL, WPIDS, JAPIO, search terms: REST, Neuron restrictive Silencer Factor, NRSF, negative regulators of neurogenesis.